Statistical Algorithms for Land/Forest Cover Change Detection Using Remote Sensing Data

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy in Engineering

by

Asim Anees

BSc. Electrical Engineering, MSc. Vision and Robotics (Erasmus Mundus, VIBOT)

School of Engineering And ICT

University of Tasmania

February, 2016

Declaration of Originality

"This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright."

Signatures: Candidate Primary Supervisor

Authority of Access

This thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.

Signatures: Candidate Primary Supervisor

Declaration of the Copyrights for the Published Chapters

"The publishers of the papers comprising Chapters 2 to/and 5 hold the copyrights for that content, and access to the material should be sought from the respective journals. The remaining non published content of the thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968."

Signatures: Candidate Primary Supervisor

Date: 11. 02. 2016

Statement of Co-Authorship

(for inclusion in the thesis)

The following people and institutions contributed to the publication of work undertaken as part of this thesis:

Candidate Name and School = **Asim Anees** (School of Engineering and ICT)
Author 2 Name and institution = **Jagannath Aryal** (School of Land and Food)
Author 3 Name and institution = **Malgorzata O'Reilly** (School of Physical Sciences, UTAS)
Author 4 Name and institution = **Timothy Gale** (School of Engineering and ICT, UTAS)
Author 5 Name and institution = **Tim Wardlaw** (Forestry Tasmania)

Authorship details:

Paper 1: Near-Rea/ Time Detection of Beetle Infestation in Pine Forests Using MODIS
Located in chapter 02
Candidate was the primary author.
Author 2 assisted on conceptual/experimental design, refinement and presentation.

Paper 2: A Statistical Framework for Near-Real Time Detection of Beetle Infestation in Pine Forests Using MODIS Data
Located in chapter 03
Candidate was the primary author.
Author 2 assisted on conceptual/experimental design, refinement and presentation.

Paper 3: A Relative Density Ratio Based Framework for Detection of Land Cover Changes in MOD/S NOVI Time-Series
Located in chapter 04
Candidate was the primary author.
Author 2, Author 3 and Author 4 assisted on conceptual/experimental design, refinement and presentation.

Paper 4: A Robust Multi-Kernel Change Detection Framework for Detecting Leaf Beetle Defolitions in Eucalyptus Nitens and E Globulus plantations Using Landsat 7 ETM+ Data
Located in chapter 05
Candidate was the primary author.
Author 2, Author 3, Author 4 and Author 5 assisted on conceptual/experimental design, refinement and presentation.

We the undersigned agree with the above stated "proportion of work undertaken" for each of the above published (or submitted) peer-reviewed manuscripts contributing to this thesis:

Signed:

Dr. Jagannath Aryal
Primary Supervisor
School of Land and Food
University of Tasmania

Prof. Andrew Chan
Head of School
School of Engineering and ICT
University of Tasmania Date:

SUMMARY

Land cover changes significantly affect climate, hydrology, bio-diversity, socio-economic stability and food security. Some of these changes being studied in remote sensing discipline include, but are not limited to, anthropogenic changes e.g. clear-cutting of forests for human settlements, and beetle/insect infestations in the forests. Beetle/insect infestations cause considerable damage to the forests resulting in tree mortality on large scale which provides fuel for fires and wastes valuable wood. Therefore, early detection of such changes is often desired by the authorities in order to carry out timely actions to mitigate them. However, manual monitoring using high resolution photography or field surveys can become very difficult and time consuming or even infeasible because such changes cover very large areas. This necessitates development of automated remote sensing algorithms which can monitor large areas with minimal human intervention. Several land cover change detection algorithms exist in literature which utilize remotely sensed imagery captured by different satellites. However, there are a very few studies which detect such changes in near-real time manner. Furthermore, there is still a room for improvement in the detection accuracy, detection delays and computational complexity of such algorithms. This thesis utilizes coarse (500 m) and moderate (30 m) spatial resolution satellite imagery (MODIS and Landsat 7 ETM+, respectively) and proposes four statistical algorithms for detection of land cover changes with significant improvements. The first algorithm (published in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing) is a supervised technique developed for near-real time detection of beetle infestations in pine forests of North America (British Columbia and Colorado). It models the hyper-temporal multi-spectral MODIS Vegetation Index (VI) time series with a triply modulated cosine function using a sliding window non-linear least squares and applies a change metric based on log-likelihood ratios to the trend parameter time-series of the fitted model, instead of the raw vegetation index. Significant improvement, in the

detection accuracy with reduced detection delays, was achieved with this first published algorithm. The second algorithm (published in IEEE Geoscience and Remote Sensing Letters) is unsupervised and makes use of properties of Martingale Central Limit Theorem (MCLT) in the change metric derived from the parameter time series, in order to avoid threshold tuning while detecting beetle infestation in MODIS vegetation index time series. The third algorithm (published in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing) avoids the Gaussian distribution based change metrics, one of the limitations of the existing methods, and improves the change detection accuracy and detection delays significantly by using assumption free, directly estimated relative density-ratio based Repeated Sequential Probability Ratio Test (RSPRT) as test statistic on the parameter time series. The fourth algorithm (submitted to ISPRS Journal of Photogrammetry and Remote Sensing) is a supervised technique that utilizes bi-temporal multi-spectral Landsat 7 ETM+ imagery and models class posterior probabilities of the change and no change classes non-parametrically, using a linear combination of a large number of Gaussian Kernels through Least Squares Probabilistic Classifier (LSPC) formulation. This helps in avoiding the Gaussian assumption about the data, which is a major drawback of the traditional Bayesian Classifiers e.g. Maximum Likelihood Classifier (MLC) and Naive Bayes (NB). Another popular classifier which avoids this limitation is Support Vector Machine (SVM) but its drawback is that it assigns only class labels to the test samples and does not provide probabilistic class-memberships. Moreover, it is a binary classifier in its original formulation and needs different strategies to be adopted in order to use it for multi-class problems and also to make it probabilistic. The LSPC based framework, on the other hand, is non-parametric as well as capable of assigning degree of class-membership (probabilistic) to test samples and handling multi-class problems in its original formulation. Its application to leaf beetle infestation problem in north-eastern Tasmanian Eucalyptus plantations suggested improvement in detection accuracy. These

methods have been compared with their respective counterparts in the literature in order to demonstrate their effectiveness.

The contribution of this thesis is four folds: (i) advancing our understanding of land/forest cover change detection using quantitative/statistical approaches, (ii) proposition of reliable statistical approaches for near-real time land/forest cover change detection in hyper-temporal coarse spatial resolution MODIS data, with especial emphasis on changes due to beetle infestations in pine forests, (iii) efficient threshold selection in complex scenario of near-real time change detection, when an optimal trade off between more than 2 performance indices is needed, and (iv) proposition of a non-parametric approach for detecting land/forest cover changes (bark beetle problem in north-eastern Tasmanian Eucalyptus plantations) in bi-temporal Landsat 7 EMT+ data.

This thesis is dedicated to:

God Almighty, for all the countless blessings and opportunities that he has bestowed upon me; My late mother, and my father for all their prayers and wholehearted sacrifices full of love and devotion, they made in raising me and enabling me to achieve success in my life; My Late grandfather, for his wholehearted hard work with me during my early school years, which made me confident about my studies, and brought a sense of healthy academic competition in my nature, which always motivated me to keep going and work hard for every degree in my educational career. Without all of these contributions from these great personalities, I would probably not have been able to come this far in my educational career with this much success.

You were born with wings, why prefer to crawl through life? -Rumi

ACKNOWLEDGEMENT

The author would like to thank the following people for their help and valuable advices during this thesis:

- Tasmanian Graduate Research Scholarship (TGRS) and University of Tasmania, for funding this PhD thesis.
- My supervisors, Jagannath Aryal, Małgorzata O'Reilly and Timothy Gale, for their support.
- Prof. J.C. Olivier (School of Engg. and ICT UTAS), for leading this project initially.
- Klyenhans Waldo (University of Pretoria), for his helpful comments, in the early stages of this thesis.
- Brian Salmon (School of Engg. and ICT UTAS), for his valuable comments and discussion during this thesis.
- Irfan Akhter Iqbal and Faheem Iqbal (School of Land and Food), for providing help in using ArcGIS software.
- Forestry Tasmania, in particular Tim Wardlaw, for providing the ground truth data for leaf beetle damage in north-eastern Tasmania.
- Mary Dixon (School of Engg. and ICT UTAS), for her administrative support in school of Engineering, throughout this thesis.

Contents

List of Figures

List of Tables

1 Introduction

Overview

Remote sensing is a process, in which information about the object under observation is obtained distantly i.e. without being in close contact with it [1]. Many remote sensors are being used to obtain data about distant objects, in order to gather useful information about them [2–5], e.g. sensors installed on satellites for monitoring atmospheric temperature and weather, telescopes used in astronomy to get information about distant planets and galaxies, radar to monitor motion of the airplanes in its range etc. It has been given immense importance over the past few decades due to its usefulness in monitoring different phenomena on the Earth, other planets and galaxies, in automated fashion, which would not have been possible otherwise. One of its important applications is monitoring of environmental changes, e.g. urban expansions, forest clear cutting, deforestation, pest infestations etc. This thesis will investigate the problem of monitoring and detection of changes in forest cover, in particular due to beetle infestations.

The aim of this chapter is to highlight the motivation behind using remote sensing for detecting insect infestations and deforestation. Section 1.1 presents definitions of some important terminologies which will be used throughout this thesis.Section 1.2 highlights the problem statement of this thesis. Section 1.3 sets the objectives to be achieved by this

thesis while approaching the problem under consideration. Section 1.4 explains different approaches used to achieve the thesis objectives. Section 1.5 gives an overview of the study, including the study areas and their maps, and the scope of the proposed methods. Finally, Section 1.6 presents the outline of rest of this thesis.

1.1 Some Definitions

Before presenting the problem statement in 1.2 and the proposed solutions, it is important to define, certain terminologies that are going to be used throughout this thesis, upfront in order to make this thesis readable and understandable.

1.1.1 Robustness

Robustness is a very well known terminology being used in engineering community, and is an important aspect while evaluating performance of different statistical algorithms. Statistical algorithms are normally based on implicit or explicit assumptions about the underlying model or distributions of the data under consideration. More often than not, these assumptions are not true and the true underlying distribution may deviate from the assumed model. In such cases, an algorithm needs to be able to handle these errors and avoid resulting in catastrophe by sacrificing a fraction of its efficiency, if need be. If an algorithm/system is able to achieve this, it called a *robust* system [6, 7]. In short, robustness demands a system/algorithm to possess the following features [7]:

- **Efficiency:** It should have near-optimal efficiency at the assumed model.
- **Stability:** It should be stable and relatively invariant to the small deviations from the model assumptions. Small model deviations should cause only small or slight impairments.

- **Avoid breakdown:** It should be able to avoid disaster (breakdown) in case of somewhat larger deviations from the model.

1.1.2 Near-real Time

Real Time is a well known terminology in automated operations which is used for characterising the time delay between the instances of the arrival of the data samples and their processing . If an automated operation or program is able to handle a process within the actual time taken by the process, without any delays, it is said to be *real time*. For example, if an automated program is performing a certain set of operations on the samples of the data being generated sequentially, and it is able to perform all the operations on the current sample before the next sample arrives (on the go) without stacking or buffering them, then it is said to be real time. In this thesis, the concept of real time has been used in a different sense. Here, rather than the time taken by the automated process, it is related to how quickly, in terms of data points, an algorithm can detect a change after its real instance of beginning, and whether it needs future data points in order to process and declare a change at the current data point (instance) or not. In other words, it is related to the delay between when (at what time point in the time series data or at what data sample) the change started to occur according to the ground truth data and when (at what time point in the time series data or at what data sample) it gets detected by the algorithm, and dependence of the algorithm on the current and past data points. Since the algorithms used in this thesis are dealing with slow changes in the forests which occur over a certain period of time (months) and take many data samples from the point in time when they start to occur till the point when they become detectable, it is nearly impossible to detect them exactly at the data samples which correspond to the instance of their occurrence in the time series. Therefore, a milder terminology *near-real time* has been used which means “not exactly real-time but as close as possible to the real instance of occurrence of

the change". Hence, the effort here is to detect the changes in the **time series data**, as close to the real instance of occurrence as possible, and to reduce the delay between the real instance of occurrence and the instance when it gets detected by the algorithm.

1.1.3 Labelling Error

The term *Labelling Error* has been used for such samples in the ground truth data, which were in fact *no-change* but labelled as *change* by the surveyors or vice-versa.. This can happen due to human errors while labelling the ground truth data.

1.1.4 True Positives

True Positive (TP) is the percentage of the test samples that are declared as *change* by the change detection algorithm, provided that they were labelled as *change* in the ground truth data.

1.1.5 True Negatives

True Negative (TN) is the percentage of the test samples that are declared as *no-change* by the change detection algorithm, provided that they were labelled as *no-change* in the ground truth data.

1.1.6 False Positives

False Positive (FP) is the percentage of the test samples that are declared as *change* by the change detection algorithm, provided that they were labelled as *no-change* in the ground truth data.

1.1.7 False Negatives

False Negative (FN) is the percentage of the test samples that are declared as *no-change* by the change detection algorithm, provided that they were labelled as *change* in the ground truth data.

1.2 Problem Statement

Deforestation and change in biomass of the Earth's forests have many implications for all those living beings who are dependent on the forests. It could be argued that we are all dependent on the forests as they are one of the main carbon sinks and provide wood to humankind, as well as a habitat for many endangered species [8]. Apart from carbon emission issues, deforestation and tree mortality is correlated with forest fires (in some parts of Canada and United States) which has both economic and environmental implications [9]. One major cause of tree mortality and the associated changes in biomass of North American forests, are beetle infestations [10–14]. This results in high risks of forest fires, increased water runoff and turbidity, loss of ecosystems services, and less absorption of the atmospheric carbon, or effectively more carbon emission, which in turn contributes to the global warming by greenhouse effect and has adverse effects on human health [11].

The beetles pass through four stages in their life cycle, namely: egg, larva, pupa and adult. It takes almost a year for them to complete this life cycle, but they can take longer than this, especially at high elevations because of the cool summer temperatures. They normally attack aged and large pine trees, which have more food and thick phloem. When they attack a part of forest, the unmated female beetles produce pheromones which attract more beetles and results in a massive attack. They reside under the bark of the pine trees and make egg galleries. These eggs then turn into larvae which feed on phloem and make

Figure 1.1: An example of Mountain Pine Beetle infestation in pine forests of British Columbia, Canada. Picture credits: Simon Fraser University (https://flic.kr/p/e8SuXx).

more galleries. After 10 months, they turn into pupae and then adults [15–17].

The beetles carry blue-staining fungi with their bodies and heads, which they spread under the bark of the infested trees. These fungi disturb the food and water supply, and reduce tree's flow of pitch. This aids the beetles to overcome the tree easily. These attacked trees then slowly turn yellow, then red and then eventually gray when they die [14]. An example of how the infested trees look like, is shown in Figure 1.1. After the tree mortality, the adults appear from the hole in the dead trees, multiplied in number as compared to their number at the time when they attacked. These beetles then head towards another nearby part of the forest [15–17]. The beetle life cycle and different stages of infestation are shown in Figure 1.2.

Effective monitoring of the forests can help the authorities to control the tree mortality by plotting timely strategies against expected occurrence of the factors causing it, e.g. beetles infestations. This could reduce the rate of global warming and wastage of expensive wood as well. However, monitoring of huge forests with human eye is not feasible. Therefore, an efficient system, that can monitor large areas automatically and detect the patterns of

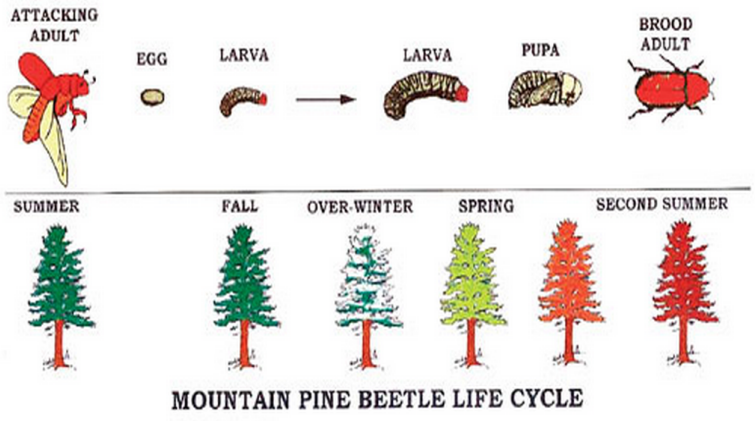

Figure 1.2: Mountain Pine Beetle life cycle and infestation stages. Picture credits: (http://www.schulhofftlc.com/insect-control/mountain-pine-beetle/).

possible beetles infestations, is required. This research is about investigating and designing robust automated techniques which use coarse or moderate resolution satellite data, e.g. MODIS (MODerate resolution Imaging Spectroradiometer) and Landsat 7 ETM+, to detect land/forest cover change, with special emphasis on tree mortality due to beetle infestations.

1.3 Objective of This Thesis

As mentioned earlier, deforestation and tree mortality, due to beetle infestations, is one of the major problems in the pine forests of the North America [10, 11]. Due to the socio-economic and environmental implications of forest mortality, beetle infestations need to be detected early so that the authorities can take timely actions to mitigate them and take precautionary measures to prevent the rest of the healthy forest from their attack. However, these forests are spread over millions of acres, and hence are too difficult to be surveyed manually or with naked eye and reported in time to the decision making authorities.

Therefore, an efficient early alarm system, that can monitor large areas automatically and detect the patterns of possible beetle infestations, will prove to be helpful in this regard [18–26].

Automated frameworks using hyper-temporal, coarse resolution satellite data, for land/forest cover change detection at regional or global scales, has been highly desired [27]. However, this goal of environmental remote sensing has been elusive and described by some as the "holy grail" of remote sensing [28]. The recent improvements in remote sensing, computational technologies and the availability of time-series analysis, patter recognition and statistical techniques have opened a new area of research about monitoring Earth systems remotely and in automatic manner. Many satellites are operating for dedicated purposes e.g. monitoring of weather, water, Earth etc. These satellites detect different processes in the atmosphere periodically, and store the image data in their databases. Some of this data is available free of cost and the rest with some cost, on world wide web (www). Since these data have been available, the researchers in remotes sensing community have made considerable contributions by analyzing this data using existing mathematical techniques [2–5, 29]. These contributions have revealed that monitoring of large areas (of the order of thousands of square kilometers) is possible to achieve, in much quicker and automated manner, by applying existing statistical time series analysis and pattern recognition techniques on the remote sensing data. So, the main objective of this thesis can be phrased as follows.

Main Objective: To improve performance of, and propose new statistical frameworks for, automated land/forest cover change detection using remotely sensed satellite data.

Remote sensing literature contains several land/forest cover change detection methods. Many articles that summarise and review these methods exist in the literature [2–5, 30]. The majority of the existing methods are based on image differencing, post-classification comparison and change trajectories of multi-date high resolution data [31–38]. In most

cases, these methods only consider two images for change detection, effectively trying to detect areas of change from one image to the next [2–5, 30]. The methods which can detect land/forest cover change in near-real time, are limited. Although a very few methods exist [39–41], they may not be necessarily optimal for beetle infestation detection in forest cover, as no single remote sensing change detection method can be assumed to be optimal for all the change detection problems [3, 4, 30]. Furthermore, near-real time change detection methods which avoid threshold tuning, based on labelled data, and also are free of parametric model assumptions, may prove more practical and robust. In case of bi-temporal remote sensing change detection, normally change vectors between two images are calculated and then classified as change or no-change. The popular classifiers used are either based on parametric models (assumption about the underlying distribution of the data) or does not have capabilities of handling multi-class problems and assigning class-membership to test samples. A classifier that can perform comparably to non-parametric classifiers and also have capabilities of handling multi-class problems and assigning class membership to test samples, may prove helpful in a wide range of remote sensing applications, apart from change detection. Therefore, the main objective of the this thesis, given above, can be split into sub-objectives as:

- To propose near-real time (if possible) framework, with improved accuracy and detection delay, for detection of land/forest cover changes, especially due to beetle infestations.
- To further make the framework more suitable for practical scenarios, by enabling it not to require labelled data for its threshold tuning.
- To propose an effective algorithm that can tune the threshold in complex near-real time scenarios, when three performance indices, instead of two as in normal cases, have to be taken into account.

- To make the proposed framework more robust for practical scenarios when the underlying data distribution does not always follow a specific parametric model, and may change from case to case.

- To evaluate new method, for remote sensing applications, that is non-parametric (free of model assumptions) as well as can handle multi-class problems and assign probabilistic class-membership to test samples.

1.4 Proposed Solution

This thesis proposes solutions to achieve the objectives, mentioned in Section 1.3, at different stages i.e. through proposed four change detection methods, while focusing on the main objective of proposing automated remote sensing methods for detecting land/forest cover changes. A brief overview of these solutions is given below and their details will be discussed in Chapters 2 through 5.

The objective of improved automated detection of gradual forest cover changes, especially beetle infestations, in near-real time using remote sensing data, is addressed in Chapter 2. As mentioned before, the remote sensing literature lacks such methods which utilise hyper-temporal satellite data to detect such changes in near-real time. Even if a few such methods exist [39–41], no method is optimal for every change detection problem, and may perform poorly in some cases [3, 4, 30]. Considering these facts, the proposed solution is based on hyper-temporal MODerate resolution Imaging Spectroradiometer (MODIS) 500 m, 8-days composite imagery. The hypothesis behind this solution is that the sensor response of the forest, after infestation, must be significantly different than when it is healthy [42, 43]. First, the MODIS vegetation index time series, of the pixel under consideration, is extracted, with at least 5 years (230 time points) of normal (unaffected or

no-change) observations in the history part. This time series is then modelled with a triply modulated cosine model [43, 44], using Non-linear Least Squares (NLS) over a moving window [42], which derives parameter time series. At every position of the window, the NLS fitting of the model produces parameters values for that point. The log likelihood ratio, between this parameter value falling inside or outside the distribution of the history (no-change) observations, is then compared with a threshold tuned on the training data, in order to declare it a change or no-change. The NLS fitting of the model handles the noise well [42] and reduces its effect in the parameter time series, hence improves the near-real time performance over the method which derive test statistics from the noise time series between the observations and modelled values [39]. The window helps in identifying the actual location of the change in the time series and makes the method near-real time. Since different vegetation indices are based on different bands and respond differently to vegetation dynamics, combining them with the proposed framework further improved the results. This method works well but it has two limitations: 1) it is supervised i.e. it needs threshold tuning and 2): the test statistics are based on Gaussian assumption which is seldom true in practical cases. Addressing these two limitations are the next objectives (sub-tasks).

In order to improve the NLS-based method further and make it suitable for more practical scenarios, a solution that does not require threshold tuning based on labelled data, is proposed in Chapter 3. In this solution, the hypothesis remains the same but the test statistics change. The normalised parameter time series, derived by fitting the triply modulated cosine function over a moving window using NLS, is modelled as a Martingale Sequence and the test statistic is derived based on the Martingale Central Limit Theorem (MCLT) [45, 46]. Once the MCLT is invoked, the threshold can be tuned with the help of standard normal distribution statistic (z-score) [47]. Experiments demonstrate the reliability of this solution.

To address the objective of avoiding Gaussian or other parametric assumptions about the data, while deriving test statistic, a solution is proposed in Chapter 4. The basic hypothesis, about the sensor response of the forest/vegetation, remains the same. The change in the framework is recommended while calculating the test statistic. Instead of using Gaussian assumption based test statistic e.g. traditional likelihood ratios that are based on parametric model assumption about the individual densities, it is recommended that relative density ratios [48, 49] that are estimated directly from the training data, using non-parametric kernel model, could be used. To this end, traditional Repeated Sequential Probability Ratios (RSPRT)/CUSUM statistics [50, 51], which use traditional log likelihood ratios, can be improved by using the relative density ratios. This solution was demonstrated by evaluating these new RSPRT/CUSUM test statistics, derived from the parameter time series of the triply modulated cosine model, which is fitted to the vegetation index time series using NLS or the Extended Kalman Filter (EKF) [51]. The experiments on simulated, synthetic and real world beetle infestation data, supports the proposed solution.

The objective of finding an efficient and automated way of tuning the threshold, on training data in case of complex scenario of near-real time, was also addressed in Chapter 4. The solution proposed here considers a more complex cost function based on three performance indices namely, False Positives (FP), False Negatives (FN) and Mean Detection Delays (MDD), than in normal cases where only FP and FN are considered [52]. Finding an optimal trade off between three performance indices, by manual observation, becomes hard or even impossible, which makes the cross validation experiments difficult. Therefore, the proposed solution argues that minimising the designed cost function using Genetic Algorithm (GA), can find an optimal trade off between the three performance indices automatically, which corresponds to the optimal threshold.

Finally, the objective of reliable detection of leaf beetle infestation in eucalyptus plantations of north-eastern Tasmania, is addressed using Landsat 7 ETM+ data. The choice of using Landsat 7 ETM+ was made because the leaf beetle infested study area was not large enough to provide sufficient number of MODIS pixel for a reliable analysis. The proposed solution is bi-temporal and supervised in nature, which first finds change vectors (by image differencing based on raw bands or vegetation indices) and then uses a supervised classifier to classify the test samples into change or no-change classes. One of the aims, while addressing this objective, was to devise a method that is independent of any parametric model assumptions about the data and immune to changes in the distribution of the data, which is a major limitation of the well known Bayesian classifiers e.g. Maximum Likelihood Classifier (MLC) and Naive Bayes (NB) classifier [53–56], and also probabilistic in nature with multi-class capabilities. This has been addressed by using non-parametric multi-kernel based LSPC formulation [57, 58], instead of model based classifiers e.g. MLC and NB, in the proposed change detection framework. The effectiveness of the proposed approach has been demonstrated through experiments and its comparison with the NB and Support Vector Machine (SVM) [59].

1.5 Scope of This Work

The type of changes targeted in this thesis were those caused by bark beetles in pine forests and leaf beetles in Eucalyptus plantations. The ground truth data for bark beetle infestations in pine forests, was collected from Colorado (USA), Utah (USA) and British Columbia (Canada) because the survey maps showing damaged areas were maintained by the forest health and pest management authorities in both these regions. Since these infestations were very widely spread, coarse spatial resolution satellite data was a reasonable option to be used for detecting such changes. The advantage of coarse spatial resolution satellite data was that it could be available with high temporal resolution, hence provid-

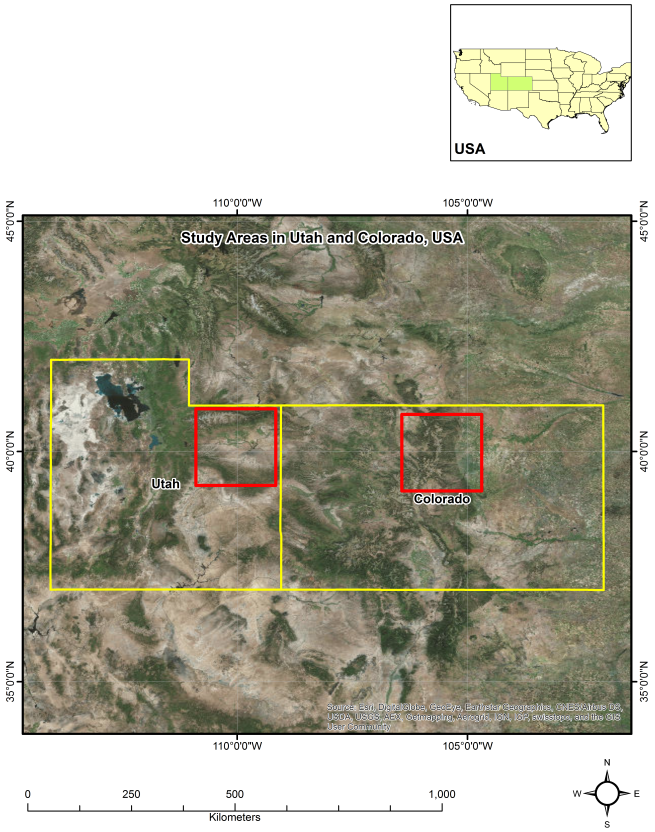

Figure 1.3: Map showing study regions (Colorado and Utah) in USA, used in Chapters 2, 3 and 4, for near-real time change detection using MODIS time series data.

ing opportunities for development of time series analysis based methods which can detect such changes with near-real time capabilities. Therefore, MODerate resolution Imaging Spectroradiometer (MODIS) 8-day composite time series data with 500m spatial resolution was used for these regions. This data was used in the development of the near-real time methods presented in Chapters 2, 3 and 4. These study regions have been shown in the maps presented in Figures 1.3 and 1.4. The details on data preparation are given in Chapter 2.

The ground truth data for leaf beetle infestations was collected from eucalyptus plantations, maintained by forest department, in north-eastern Tasmania, Australia (Figure 1.5). This area was used in the development of a change detection method based on image

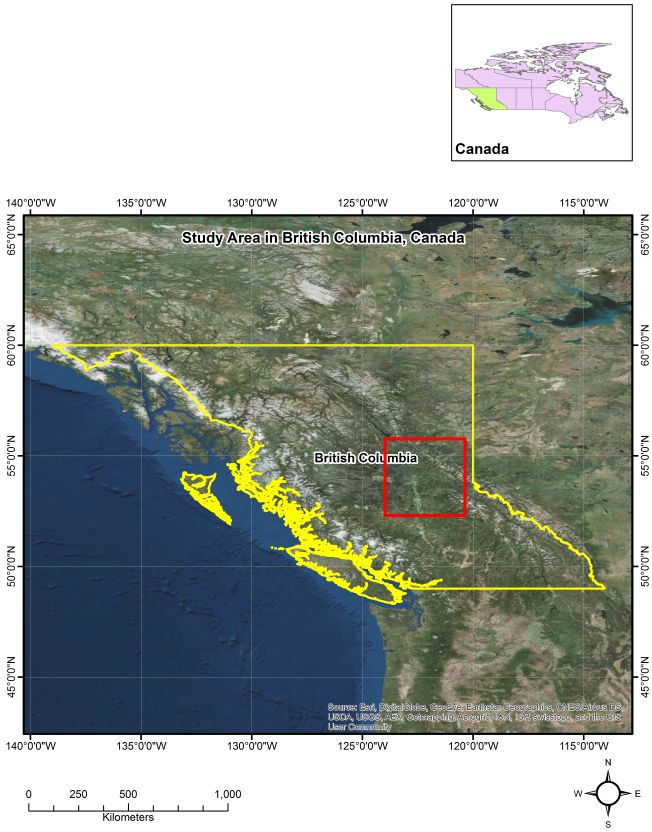

Figure 1.4: Map showing study region (in British Columbia) in Canada, used in Chapters 2, 3 and 4, for near-real time change detection using MODIS time series data.

differencing as an initial step. Since, these plantations were limited in size and area, the changes recoded here were on much smaller scales, spatial resolution of MODIS data was not appropriate for this analysis. Therefore, a novel change detection method based on Landsat 7 ETM+ 30m imagery was proposed to tackle this problem.

Although the methods presented here have been designed and tested for beetle infestation problem in pine forests and eucalyptus plantations, remote sensing literature is rich with other applications of the statistical change detection algorithms [60–80]. Some examples of different applications of change detection methods from recent literature are: 1) Crop Changes: Remote sensing change detection algorithms can be applied to detect changes in croplands e.g. from one crop type to another crop or land cover type or ex-

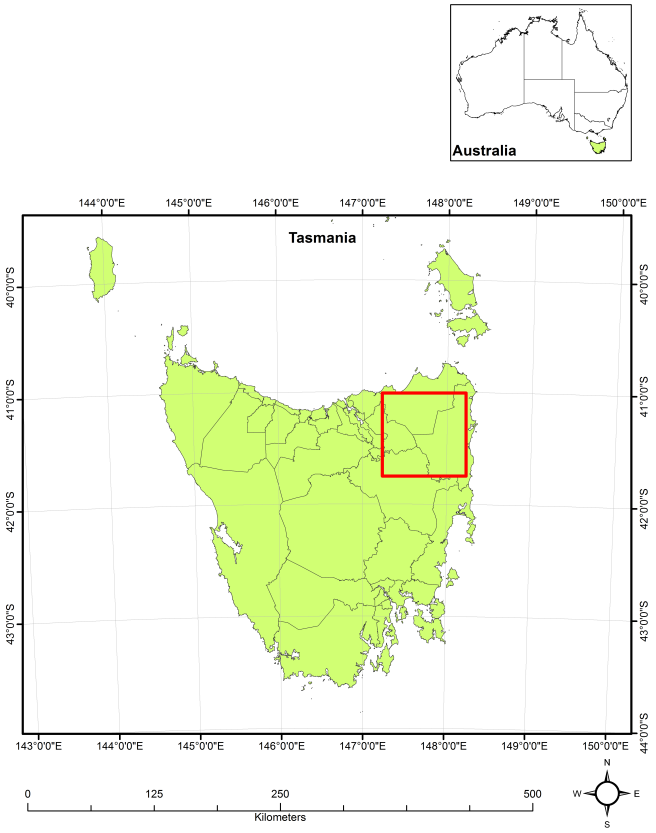

Figure 1.5: Map showing study region (in Tasmania) in Australia, used in Chapters 5, for multi-temporal change detection using Landsat 7 ETM+ data.

pansion/shrinking of the land area covered by a certain type of crops [66, 67], 2) Urban Changes: Remote sensing change detection algorithms are also used to detect new settlements, construction of new buildings/roads/facilities etc. in the urban environments, e.g. [75–77], 3) Building Changes: Changes in structures or demolition of the existing buildings can also be monitored using remote sensing change detection algorithms, e.g. [64, 65], 4) Hydrological Changes: Monitoring of hydrological changes in soil e.g. monitoring of changes in wetland, soil moisture, soil state, salinity, erosion and deposition in coastal areas, can also be one of the applications of remote sensing change detection methods, e.g. [60, 61, 79], 5) Water Changes: Such methods can also be applied for detecting changes in water levels or lakes and different water reservoirs, e.g [80], 6) Flood

Changes: Changes induced by floods on different land cover types or flood itself can also be detected and monitored using such techniques, e,g, [71–73], 7) Fire Changes: Remote sensing change detections can be applied to detect active forest fires or changes induced by forest fires, e.g. [69, 70], and 7) Deforestation: Detection of deforestation, caused by logging and human settlements, is also a potential application of remote sensing change detection methods, e.g. [74, 81, 82]. Since the methods presented in this thesis are statistical in nature and rely on statistics of different features of the data, their scope is not limited to forest cover change detection only, and can be adapted to many different change detection applications mentioned above.

1.6 Thesis Outline

- Chapter 2 presents the first method proposed in this thesis, that uses MODIS hyper-temporal time series data for near-real time detection of beetle infestations in pine forests of British Columbia, Canada and Colorado, USA. Two methods are investigated: 1) an NLS-based method that uses a triply modulated cosine function to model the vegetation index time series and then finds a change in its parameter time series, and 2) an FIR low-pass filter based method, which does not include any model fitting, hence a computationally faster and less complex. The chapter also discusses combination of these methods with several vegetation indices and finds the best combination for near-real time beetle infestation detection. In the end, the chapter presents the results of the proposed framework, and its comparison with an existing near-real time change detection method.
- Chapter 3 presents the second method proposed in this thesis, that uses MODIS hyper-temporal time series data for near-real time detection of beetle infestations

in pine forests. It discusses the limitations in case of supervised frameworks when threshold tuning is needed based on the labelled data sets, and how the MCLT-based statistics, derived from the parameter time series, helps in this regard. The chapter investigates this framework on a variety of vegetation index data sets and presents its results in the end, in order to show the promise of the proposed framework.

- Chapter 4 presents the third method proposed in this thesis, which introduces improvements in the supervised framework for near-real time change detection. It highlights a bottleneck of the traditional likelihood ratios based test statistics which are based on Gaussian or other parametric model assumptions about the underlying density of the data, and advocates replacing it by more robust relative density ratios which are estimated directly from the data using a non-parametric model. It further discusses, how the relative density ratios improve the RSPRT/CUSUM statistics which are derived from the parameter time series of the triply modulated cosine function, fitted to the data using NLS or EKF. In the end, it discusses the promise shown by the proposed framework through its implementation for near real-time detection of beetle infestation and the comparison with the existing methods that are based on the traditional likelihood ratios.

- Chapter 5 presents the fourth method proposed in this thesis, for detection of defoliation, due to leaf beetle in eucalyptus trees of north-eastern Tasmania, using Landsat 7 ETM+ imagery. The chapter highlights the limitations of the well known Bayesian classifiers e.g. MLC and NB, which is still being used in remote sensing for classification and change detection, for changes in data distribution. It further explains how this limitation can be avoided by using a non-parametric multi-kernel LSPC formulation for modeling the class posterior probabilities. Furthermore, it presents an automated model selection procedure that avoids manual selection of possibly sub-optimal model parameters through 10-fold cross-validation. In the

end, it discusses the improvements in the performance brought by the proposed framework along with the proposed model selection procedure, by comparing its results with an NB and SVM [59] based methods, while detecting leaf beetle defoliation in eucalyptus trees of north-eastern Tasmania, using Landsat 7 ETM+ multi-date imagery.

- Chapter 6 presents key conclusions, as well as, future research ideas which can further enhance the capability of remote sensing change detection algorithms.

Bibliography

[1] T. Lillesand and R. Kiefer, *Remote Sensing and Image Interpretation*. New York, NY: John Wiley and Sons, 4th ed., 2000.

[2] P. C. Smits and A. Annoni, "Toward specification-driven change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 3, pp. 1484–1488, 2000.

[3] D. Lu, P. Mausel, E. Brondizio, and E. Moran, "Change detection techniques," *International journal of remote sensing*, vol. 25, no. 12, pp. 2365–2401, 2004.

[4] P. Coppin, I. Jonckheere, K. Nackaerts, B. Muys, and E. Lambin, "Digital change detection methods in ecosystem monitoring: a review," *International journal of remote sensing*, vol. 25, no. 9, pp. 1565–1596, 2004.

[5] A. Singh, "Review article digital change detection techniques using remotely-sensed data," *International journal of remote sensing*, vol. 10, no. 6, pp. 989–1003, 1989.

[6] F. J. Anscombe, "Rejection of outliers," *Technometrics*, vol. 2, no. 2, pp. 123–146, 1960.

[7] P. J. Huber and E. Ronchetti, "Robust statisticswiley," *New York*, 1981.

[8] "Forests: Climate change, biodiversity and land degradation," tech. rep., Joint Liaison Group of the Rio Conventions.

[9] R. Watson, I. Noble, B. Bolin, N. Ravindranath, D. Verardo, and D. Dokken, "Land Use, Land-Use Change and Forestry," *Cambridge University Press, UK*, p. 375, 2000.

[10] R. R. Schneider, M. C. Latham, B. Stelfox, D. Farr, and S. Boutin, "Effects of a severe mountain pine beetle epidemic in western alberta, canada under two forest management scenarios," *International Journal of Forestry Research*, vol. 2010, no. 417595, p. 7, 2010.

[11] S. Embrey, J. V. Remais, and H. J., "Climate change and ecosystem disruption: the health impacts of the north american rocky mountain pine beetle infestation," *American Journal of Public Health*, vol. 5, pp. 818–827, May 2012.

[12] A. J. Meddens, J. A. Hicke, and L. A. Vierling, "Evaluating the potential of multispectral imagery to map multiple stages of tree mortality," *Remote Sensing of Environment*, vol. 115, pp. 1632–1642, Jul 2011.

[13] A. J. Meddens, J. A. Hicke, L. A. Vierling, and A. T. Hudak, "Evaluating methods to detect bark beetle-caused tree mortality using single-date and multi-date landsat imagery," *Remote Sensing of Environment*, vol. 132, pp. 49–58, May 2013.

[14] M. A. Wulder, C. C. Dymond, J. C. White, D. G. Leckie, and A. L. Carroll, "Surveying mountain pine beetle damage of forests: A review of remote sensing opportunities," *Forest Ecology and Management*, vol. 221, no. 1, pp. 27–41, 2006.

[15] G. D. Amman and W. E. Cole, "Mountain pine beetle dynamics in lodgepole pine forests. part ii: population dynamics.," *Gen. Tech. Rep. INT-146. Ogden,*

UT:U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, 1983.

[16] D. Leatherman, I. Aguayo, and T. Mehall, "Mountain pine beetle," *Colorado State Forest Service*, vol. 5.528, 2012.

[17] C. F. Service, "http://www.pc.gc.ca/eng/docs/v-g/dpp-mpb/index.aspx," *Parks Canada*, 2009.

[18] R. H. Fraser and R. Latifovic, "Mapping insect-induced tree defoliation and mortality using coarse spatial resolution satellite imagery," *International Journal of Remote Sensing*, vol. 26, pp. 193–200, Jan 2005.

[19] S. Franklin, M. Wulder, R. Skakun, and A. Carroll, "Mountain pine beetle red-attack forest damage classification using stratified landsat tm data in british columbia, canada," *Photogrammetric Engineering & Remote Sensing*, vol. 69, pp. 283–288, Mar 2003.

[20] N. R. Goodwin, N. C. Coops, M. A. Wulder, S. Gillanders, T. A. Schroeder, and T. Nelson, "Estimation of insect infestation dynamics using a temporal sequence of landsat data," *Remote Sensing of Environment*, vol. 112, pp. 3680–3689, Sep 2008.

[21] N. R. Goodwin, S. Magnussen, N. C. Coops, and M. A. Wulder, "Curve fitting of time-series landsat imagery for characterizing a mountain pine beetle infestation," *International Journal of Remote Sensing*, vol. 31, pp. 3263–3271, Jun 2010.

[22] R. S. Skakun, M. A. Wulder, and S. E. Franklin, "Sensitivity of the thematic mapper enhanced wetness difference index to detect mountain pine beetle red-attack damage," *Remote Sens. Environ.*, vol. 86, p. 433, Aug 2003.

[23] J. White, M. Wulder, D. Brooks, R. Reich, and R. Wheate, "Detection of

red attack stage mountain pine beetle infestation with high spatial resolution satellite imagery," *Remote Sensing of Environment*, vol. 96, pp. 340–351, Jun 2005.

[24] J. White, W. M.A., and D. Grills, "Detecting and mapping mountain pine beetle red-attack damage with spot-5 10-m multispectral imagery," *BC Journal of Ecosystems and Management*, vol. 7, no. 2, pp. 105 – 118, 2006.

[25] J. C. White, N. C. Coops, T. Hilker, M. A. Wulder, and A. L. Carroll, "Detecting mountain pine beetle red attack damage with eo-1 hyperion moisture indices," *International Journal of Remote Sensing*, vol. 28, pp. 2111–2121, May 2007.

[26] M. Wulder, J. White, B. Bentz, M. Alvarez, and N. Coops, "Estimating the probability of mountain pine beetle red-attack damage," *Remote Sensing of Environment*, vol. 101, pp. 150–166, Mar 2006.

[27] J. R. G. Townshend and C. O. Justice, "Selecting the spatial resolution of satellite sensors required for global monitoring of land transformations," *International Journal of Remote Sensing*, vol. 9, no. 2, pp. 187–236, 1988.

[28] T. Loveland, T. Sohl, A. Stehman, S.V.and Gallant, K. Sayler, and D. Napton, "A strategy for estimating the rates of recent united states land-cover changes," *Photogrammetric Engineering & Remote Sensing*, vol. 68, no. 10, pp. 1091–1099, 2002.

[29] L. Bruzzone and D. F. Prieto, "A minimum-cost thresholding technique for unsupervised change detection," *International Journal of Remote Sensing*, vol. 21, no. 18, pp. 3539–3544, 2000.

[30] D. Lu and Q. Weng, "A survey of image classification methods and techniques

for improving classification performance," *International journal of Remote sensing*, vol. 28, no. 5, pp. 823–870, 2007.

[31] T. Celik, "A bayesian approach to unsupervised multiscale change detection in synthetic aperture radar images," *Signal processing*, vol. 90, no. 5, pp. 1471–1485, 2010.

[32] T. Celik and K. Ma, "Unsupervised change detection for satellite images using dual-tree complex wavelet transform," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 48, no. 3, pp. 1199–1210, 2010.

[33] F. Bovolo and L. Bruzzone, "A split-based approach to unsupervised change detection in large-size multitemporal images: Application to tsunami-damage assessment," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 6, pp. 1658–1670, 2007.

[34] J. S. Deng, K. Wang, Y. Hong, and J. G. Qi, "Spatio-temporal dynamics and evolution of land use change and landscape pattern in response to rapid urbanization," *Landscape and Urban Planning*, vol. 92, no. 3, pp. 187–198, 2009.

[35] C. S. JHA and N. V. M. Unni, "Digital change detection of forest conversion of a dry tropical indian forest region," *International Journal of Remote Sensing*, vol. 15, no. 13, pp. 2543–2552, 1994.

[36] J. O. Townshend, J. R. G.RG Townshend, "Spatial variability of images and the monitoring of changes in the normalized difference vegetation index," *International Journal of Remote Sensing*, vol. 16, no. 12, pp. 2187–2195, 1995.

[37] E. F. Lambin and A. H. Strahlers, "Change-vector analysis in multitemporal space: a tool to detect and categorize land-cover change processes using high temporal-resolution satellite data," *Remote Sensing of Environment*, vol. 48, no. 2, pp. 231–244, 1994.

[38] F. J. Garcia-Haro and J. Gilabert, M. A .and Melia, "Monitoring fire-affected areas using thematic mapper data," *International Journal of Remote Sensing*, vol. 22, no. 4, pp. 533–549, 2001.

[39] J. Verbesselt and M. Zeileis, A.and Herold, "Near real-time disturbance detection using satellite image time series," *Remote Sensing of Environment*, vol. 123, pp. 98–108, 2012.

[40] J. P. Spruce, S. Sader, R. E. Ryan, J. Smoot, P. Kuper, K. Ross, D. Prados, J. Russell, G. Gasser, R. McKellip, and W. Hargrove, "Assessment of modis ndvi time series data products for detecting forest defoliation by gypsy moth outbreaks," *Remote sensing of environment*, vol. 115, no. 2, pp. 427–437, 2011.

[41] Q. Xin, P. Olofsson, Z. Zhu, B. Tan, and C. E. Woodcock, "Toward near real-time monitoring of forest disturbance by fusion of modis and landsat data," *Remote Sensing of Environment*, vol. 135, pp. 234–247, 2013.

[42] A. Anees and J. Aryal, "Near-real time detection of beetle infestation in pine forests using modis data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 99, pp. 1–11, 2014.

[43] W. Kleynhans, J. C. Olivier, K. J. Wessels, B. P. Salmon, F. Van Den Bergh, and K. Steenkamp, "Detecting land cover change using an Extended Kalman Filter on MODIS NDVI time-series data," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 3, pp. 507–511, 2011.

[44] W. Kleynhans, J. C. Olivier, K. J. Wessels, F. van den Bergh, B. P. Salmon, and K. C. Steenkamp, "Improving land cover class separation using an extended kalman filter on modis ndvi time-series data," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, pp. 381–385, Apr 2010.

[45] P. Hall and C. C. Heyde, *Martingale limit theory and its application*. Academic, 1980.

[46] B. M. Brown, "Martingale central limit theorems," *The Ann. of Math. Stats.*, vol. 42, pp. 59–66, Feb 1971.

[47] D. Montgomery, *Introduction to statistical quality control*. New York, USA: Wiley, 3rd ed., 1997.

[48] S. Liu, M. Yamada, N. Collier, and M. Sugiyama, "Change-point detection in time-series data by relative density-ratio estimation," *Neural Networks*, vol. 43, no. 0, pp. 72 – 83, 2013.

[49] Y. Kawahara and M. Sugiyama, "Change-point detection in time-series data by direct density-ratio estimation," in *Proceedings of the SIAM International Conference on Data Mining, SDM 2009, April 30 - May 2, 2009, Sparks, Nevada, USA*, pp. 389–400, SIAM, 2009.

[50] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1/2, pp. pp. 100–115, 1954.

[51] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[52] J. Im, J. Rhee, J. R. Jensen, and M. E. Hodgson, "An automated binary change detection model using a calibration approach," *Remote Sensing of Environment*, vol. 106, no. 1, pp. 89 – 105, 2007.

[53] R. A. Redner and H. F. Walker, "Mixture densities, maximum likelihood and the em algorithm," *SIAM review*, vol. 26, no. 2, pp. 195–239, 1984.

[54] R. O. Duda and P. E. Hart, *Pattern recognition and scene analysis*, vol. 32. Wiley, New York, 1973.

[55] I. Rish, "An empirical study of the naive bayes classifier," tech. rep., 2001.

[56] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," *Mach. Learn.*, vol. 29, pp. 131–163, Nov. 1997.

[57] M. Sugiyama, "Superfast-trainable multi-class probabilistic classifier by least-squares posterior fitting," *IEICE Trans. Inf. & Syst.*, vol. 93, pp. 2690–2701, oct 2010.

[58] M. Sugiyama and J. Simm, "A computationally-efficient alternative to kernel logistic regression," in *IEEE International Workshop on Machine Learning for Signal Processing (MLSP), 2010*, pp. 124–129, Aug 2010.

[59] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag New York, Inc., 1995.

[60] G. P. Petropoulos, D. P. Kalivas, H. M. Griffiths, and P. P. Dimou, "Remote sensing and gis analysis for mapping spatio-temporal changes of erosion and deposition of two mediterranean river deltas: The case of the axios and ali-akmonas rivers, greece," *International Journal of Applied Earth Observation and Geoinformation*, vol. 35, pp. 217–228, 2015.

[61] B. Tian, Y.-X. Zhou, R. M. Thom, H. L. Diefenderfer, and Q. Yuan, "Detecting wetland changes in shanghai, china using {FORMOSAT} and landsat {TM} imagery," *Journal of Hydrology*, vol. 529, Part 1, pp. 1 – 10, 2015.

[62] A. P. Tewkesbury, A. J. Comber, N. J. Tate, A. Lamb, and P. F. Fisher, "A critical synthesis of remotely sensed optical image change detection techniques," *Remote Sensing of Environment*, vol. 160, pp. 1 – 14, 2015.

[63] D. Lu, G. Li, and E. Moran, "Current situation and needs of change detection techniques," *International Journal of Image and Data Fusion*, vol. 5, no. 1, pp. 13–38, 2014.

[64] X. Huang, L. Zhang, and T. Zhu, "Building change detection from multi-

temporal high-resolution remotely sensed images based on a morphological building index," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, pp. 105–115, Jan 2014.

[65] S. Nebiker, N. Lack, and M. Deuber, "Building change detection from historical aerial photographs using dense image matching and object-based image analysis," *Remote Sensing*, vol. 6, no. 9, p. 8310, 2014.

[66] J. Dong, X. Xiao, W. Kou, Y. Qin, G. Zhang, L. Li, C. Jin, Y. Zhou, J. Wang, C. Biradar, J. Liu, and B. M. III, "Tracking the dynamics of paddy rice planting area in 1986-2010 through time series landsat images and phenology-based algorithms," *Remote Sensing of Environment*, vol. 160, pp. 99 – 113, 2015.

[67] Z. Zhu and C. E. Woodcock, "Continuous change detection and classification of land cover using all available landsat data," *Remote sensing of Environment*, vol. 144, pp. 152–171, 2014.

[68] K. S. Willis, "Remote sensing change detection for ecological monitoring in united states protected areas," *Biological Conservation*, vol. 182, pp. 233 – 242, 2015.

[69] W. Schroeder, P. Oliva, L. Giglio, B. Quayle, E. Lorenz, and F. Morelli, "Active fire detection using landsat-8/oli data," *Remote Sensing of Environment*, 2015.

[70] J. T. Arnett, N. C. Coops, L. D. Daniels, and R. W. Falls, "Detecting forest damage after a low-severity fire using remote sensing at multiple scales," *International Journal of Applied Earth Observation and Geoinformation*, vol. 35, pp. 239–246, 2015.

[71] D. Lingadevaru, P. Jayakumar, *et al.*, "Flood induced land use and land cover

change detection using remote sensing and gis-a study of dhadesugur and siraguppa subwaterheds in lower tungabhadra catchment, karnataka," *International Journal of Geomatics and Geosciences*, vol. 5, no. 4, p. 573, 2015.

[72] V. Klemas, "Remote sensing of floods and flood-prone areas: An overview," *Journal of Coastal Research*, vol. 31, no. 4, pp. 1005–1013, 2014.

[73] Y. Wang, "Advances in remote sensing of flooding," *Water*, vol. 7, no. 11, pp. 6404–6410, 2015.

[74] J. Reiche, S. de Bruin, D. Hoekman, J. Verbesselt, and M. Herold, "A bayesian approach to combine landsat and alos palsar time series for near real-time deforestation detection," *Remote Sensing*, vol. 7, no. 5, p. 4973, 2015.

[75] G. Cao, Y. Li, Y. Liu, and Y. Shang, "Automatic change detection in high-resolution remote-sensing images by means of level set evolution and support vector machine classification," *International Journal of Remote Sensing*, vol. 35, no. 16, pp. 6255–6270, 2014.

[76] B. Wang, S. Choi, Y. Byun, S. Lee, and J. Choi, "Object-based change detection of very high resolution satellite imagery using the cross-sharpening of multitemporal data," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, pp. 1151–1155, May 2015.

[77] I. R. Hegazy and M. R. Kaloop, "Monitoring urban growth and land use change detection with {GIS} and remote sensing techniques in daqahlia governorate egypt," *International Journal of Sustainable Built Environment*, vol. 4, no. 1, pp. 117 – 124, 2015.

[78] Q. Wang, P. M. Atkinson, and W. Shi, "Fast subpixel mapping algorithms for subpixel resolution change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, pp. 1692–1706, April 2015.

[79] K. Rautiainen, T. Parkkinen, J. Lemmetyinen, M. Schwank, A. Wiesmann, J. Ikonen, C. Derksen, S. Davydov, A. Davydova, J. Boike, *et al.*, "Smos prototype algorithm for detecting autumn soil freezing," *Remote Sensing of Environment*, 2016.

[80] K. Rokni, A. Ahmad, A. Selamat, and S. Hazini, "Water feature extraction and change detection using multitemporal landsat imagery," *Remote Sensing*, vol. 6, no. 5, p. 4173, 2014.

[81] W. Kleynhans, B. Salmon, K. Wessels, and J. Olivier, "Rapid detection of new and expanding human settlements in the limpopo province of south africa using a spatio-temporal change detection method," *International Journal of Applied Earth Observation and Geoinformation*, vol. 40, pp. 74 – 80, 2015.

[82] R. B. Mapfumo, A. Murwira, M. Masocha, and R. Andriani, "Detection of subtle deforestation due to logging using satellite remote sensing in wet and dry savanna woodlands of southern africa," *Geocarto International*, vol. 0, no. 0, pp. 1–17, 2016.

2 Near-Real Time Detection of Beetle Infestation in Pine Forests Using MODIS Data[1]

2.1 Overview

The paper considers near-real time detection of beetle infestation in North American pine forests using MODIS 8-days 500 m data. Two methods are considered, both using a single time-series for detection of beetle infestation by analyzing the statistics of the trend component of the signal. The first method estimates the trend component of the vegetation index time-series by fitting an underlying triply modulated cosine model over a sliding window, using Nonlinear Least Squares (NLS), and the second method uses a $T-$point moving average Finite Impulse Response (FIR) filter. Both the methods perform well and show similar performance on simulated datasets. The methods are also tested on many *difference* and *ratio-indices* of a real world dataset with change and no-change examples taken from the Rocky Mountain region of the United States and of British Columbia in Canada. The results suggest that both the methods detect beetle infestation reliably in

[1]This chapter has been published as a journal article in IEEE **J**ournal of **S**elected **T**opics in **A**pplied Earth Observation and **R**emote **S**ensing (IEEE JSTARS).

almost all the vegetation index datasets. However, the model-based method (NLS-based) performs better in terms of the detection delay. Red Green Index (RGI), when used with the model-based method, provides the best trade-off between the detection delay and accuracy. Furthermore, 90%, 50% and 25% cross-validations are also performed for the threshold selection on RGI dataset, and it is shown that the selected threshold works well on the test data. In the end, it is also shown that the model-based method outperforms a recently published method for near-real time disturbance detection in MODIS data, in both accuracy and detection delay.

2.2 Introduction

Change in biomass is associated with land cover change, which occurs due to natural and human activities [1]. Automated land cover change detection using remotely sensed data is a topic of ongoing research, and has a significant body of literature [2–38]. Of specific interest in this paper is detecting gradual change in forests due to insect infestation, specifically the North American beetle [2, 20–22, 39]. The North American beetle has caused significant damage to the pine forests over the past few decades [20, 21]. Therefore, early and reliable detection of the beetle infestation is important.

Monitoring large forest using manual techniques is not feasible, and is very time consuming. There is a need to develop efficient systems, as an alternative, to monitor large areas automatically, and detect the patterns of possible beetle infestation. The choice of a satellite sensor with reasonable cost and appropriate resolution (spatial, spectral and temporal) of imagery is crucial [37]. A possible method is to make use of coarse spatial resolution images available at a high temporal rate. This has proven to be challenging, but can serve as a first step in tasking high resolution satellites for expected change events [40].

The literature contains significant number of studies on detection of beetle infestations [6, 11, 12, 24, 33–36, 41–47]. Most of these studies have utilized either high resolution

($\leq$10m spatial resolution) [2, 6, 33, 34, 48] or moderate resolution ($>$10m and $<$30m spatial resolution) sensor data [7, 11, 20–22, 35, 42, 43, 45]. Most of the high/moderate resolution sensors do not provide data at regular time intervals, hence the majority of the existing methods rely upon analysis of two images or a few images per year. This may result in suboptimal performance because of the seasonal variations in the signal received from vegetation. Therefore, the temporal resolution of the data should be high enough to differentiate real change events from the natural phenological cycles [16, 19]. MODIS provides high temporal and coarse spatial resolution data at regular time intervals i.e. daily, 8-days and 16-days composites. Several studies on land cover change detection have used MODIS data and reported impressive results [5, 16, 19, 25, 30–32]. Some of them have also investigated feasibility of MODIS data for detection of insect-induced tree/forest mortality [10, 27, 29, 49], but they have not utilized its full temporal resolution. In this paper, we propose a framework that uses MODIS 8-days composite 500m data, and utilize its full temporal resolution in monitoring beetle infestation.

It has been shown that the MODIS Normalized Difference Vegetation Index (NDVI) time-series is useful in monitoring vegetation [16, 19]. The NDVI time-series of MODIS data has a strong seasonal component, and it was proposed in [16] that the MODIS NDVI time-series be modeled as a triply modulated cosine function (2.1). Some studies [50, 51] used Fast Fourier Transformation (FFT) to analyze the spectral components of the NDVI time-series in order to detect land cover changes. However, [15] showed that the parameters of the triply modulated cosine function (2.1) provides better separation between vegetation and non-vegetation classes. The study in [16] used Extended Kalman Filter (EKF) to derive these parameters for change detection. The change metric used in [16] compares EKF-modeled parameters of (2.1) for a given pixel with those of its neighboring pixels in a 3×3 window. This method performs well when the change is a spatially-rare event, i.e. when the changed pixels in the window are a minority, and loses its accuracy for the case

of changes caused by beetles in pine forests. This is because these changes are normally widespread and involve the majority of the pixels in the window [52].

In [52], we proposed modifications in the EKF-based method [16] for detection of beetle infestation, and introduced an Non-linear Least Squares (NLS) based method for further improvements in the results. We showed that the beetle infestation can be detected well by modeling a single time-series at a time using the triply modulated cosine function (2.1) and the trend parameter (μ_k) of the function plays a key role in detecting beetle infestation. Furthermore, we showed that the statistics of the time varying trend parameter (μ_k) of (2.1), when derived by using EKF or by fitting (2.1) to the NDVI time-series using NLS, were significantly different for change and no-change periods. This difference in the statistics of the μ_k time-series was used as a tool for fast and reliable detection of beetle infestation in offline mode. It was found that the NLS-based method performed better in terms of both the accuracy and the detection delay.

This paper extends the NLS-based method, as a model-based method for near-real time detection of beetle infestation. Furthermore, a computationally simpler method based on moving average Finite Impulse Response (FIR) filter, here on referred to as FIR-based method, is also considered for near-real time change detection. Both the methods are compared first on simulated data, for sensitivity, robustness against different levels of noise, and detection delay. The results show that both the methods detect changes in the simulated data reliably, and robustly in case of different levels of signal noise. It is also shown that both the methods can detect changes in near-real time, even when the magnitude of change is very small. The proposed methods are also compared on different vegetation indices calculated from the real world dataset of change and no-change examples taken from two study regions in North America (British Columbia and Colorado). These vegetation indices are: Normalized Difference Vegetation Index (NDVI) [53], Normalized Difference Infrared Index 6 (NDII6) [10, 54], Normalized Difference Infrared Index 7

(NDII7) [10, 54], Normalized Difference Water Index (NDWI) [10, 55], Green Normalized Difference Vegetation Index (GNDVI) [56], Wide Dynamic Range Vegetation Index (WDRVI) [57], Simple Ratio (SR) [12], Infrared Simple Ratio 6 (ISR6) [12], Infrared Simple Ratio 7 (ISR7) [12] and Red Green Index (RGI) [6]. The results show that both the methods achieve similar detection accuracies, but the model-based method detects the changes faster than the FIR-based method. Among different indices, the model-based method performs the best on RGI, in terms of detection delay, whereas detection accuracies remain nearly the same for all the indices, except the NDWI. The results also suggest that the methods generally perform better with the *ratio-indices* than with the *difference-indices*. An analysis is also performed to suggest the optimal window size for the model-based method. This method is then cross-validated on the RGI data using 10%, 50% and 75% of the dataset as test data. It is shown that the model-based method achieves nearly the same accuracies on all the three cross-validations, hence suggesting that the selected threshold is robust for the test data.

Recently, a study was published in [32], which uses a season-trend function to model 16-day MODIS NDVI time-series and detects changes in test statistics derived from the difference (noise) sequence between the model and actual observations. We found this study to be the most recent and relevant to our work, in remote sensing literature. We compare our model-based method with this framework, on simulated as well as real datasets, and show that our method, which detects changes in parameter time-series, performs better in terms of both accuracy and detection delay.

The paper is organized as follows. Section 2.3 explains the proposed methods, the most recent and relevant methodology in literature, the simulation setups, and collection of the ground truth data. Section 2.4 presents the numerical results and discussions, and the last section concludes this paper.

2.3 Materials and Methods

In this section, we propose two methods for detecting changes induced by beetle infestation in pine forests. The first method is model-based and uses a parameter time-series for detecting the beetle infestation. The parameter time-series is derived by fitting a triply modulated cosine model (2.1) to the vegetation index time-series, using NLS [58, 59] over a sliding window. The second method uses a simple FIR (Finite Impulse Response) filter, implementing T-point moving averages [60].

2.3.1 Model-Based Method (Method 1)

Let Y_k represents vegetation index time-series for a single pixel which can be modeled using the triply modulated cosine function [15, 16] as:

$$Y_k = \mu_k + \alpha_k \cos(2\pi f k + \phi_k) + \varepsilon_k \,, \tag{2.1}$$

where Y_k, μ_k, α_k, ϕ_k and ε_k are the observation, trend, amplitude of the cosine term, phase and noise values, respectively, at time k, for all $k = 1, \ldots, N$. The frequency f in (2.1) depends upon the length of the seasonal cycle and the number of days between the two consecutive data acquisition points. Since, we are using MODIS 8-days composite time-series data, which has temporal resolution of 8 days and cycle length of 365 days, the value of f in our case is $\frac{8}{365}$. Stationary time-series with seasonality can be modeled using a single cosine function with constant parameters, $\mathbf{x} = [\mu, \alpha, \phi]^t$, as:

$$Y_k = \mu + \alpha \cos(2\pi f k + \phi) + \varepsilon_k \,. \tag{2.2}$$

The model in (2.2) fails to capture any non-stationarity that might be encountered within the time-series, because it has only one sinusoidal term and constant parameters. Since, these parameters are derived by taking a holistic view of the entire time-series, they are

not sufficient to cater significantly for small and short duration seasonal/trend variations. Such models are less helpful or not helpful at all in the problem of time-series change detection [61]. To cope up with such variations, the model needs to have a sufficiently greater degree of freedom/flexibility, so that it can adjust to any changes/variations in seasonal cycle or trend of the time-series. Models with sufficiently large number of sinusoidal terms and parameters, such as FFT [50, 51], provide more significant information about such variations, but sometimes these models too prove less helpful as compared to time-varying parameters-based models, such as (2.1) [15]. The advantage of the triply modulated cosine model (2.1), over FFT or model given in (2.2), is that it has three time-varying parameters, which can vary at every time point and adapt to any small, large, slow or fast variations in the time-series.

We use local fitting of the model in (2.2), over a sliding window, to derive the instantaneous parameters of (2.1). The time-series under consideration is either dominantly stationary or dominantly non-stationary over small periods. Either of these effects will be reflected more significantly in the constant parameters of the model (2.2) fitted over a window than in the parameters derived by fitting the same model over the whole time-series. Fitting (2.2) to a sliding window (overlapping windows) in continuum derives time-series of the three constant parameters, which can be considered as the time-varying parameters of model (2.1). This process of model fitting will reflect the transition of the original time-series, from stationarity to non-stationarity, in one or more parameter time-series. This is because the parameters of (2.2), derived for windows corresponding to the stationary and non-stationary parts of the time-series, will be significantly different. Moreover, the effect of every single time-point is translated into one or more parameter time-series which makes the change detection method near-real time. Parameters of (2.2) derived by fitting it globally to a time-series give no information about when the change occurred within the time-series. If the change is large and affects these constant parameters significantly,

still it requires the whole time-series to be available and works only in offline mode. We propose the following strategy for deriving the parameters of (2.1) and then detecting the beetle infestation (change).

We take a window $[k-T+1, k]$, of length T, and slide it over the vegetation index time-series. At every position of the sliding window, $k = T, T+1, \ldots, N$, the time-series covered by it can be modeled as:

$$Y_i = F(i) + \varepsilon_i, \tag{2.3}$$

where

$$F(i) \equiv F(i; \mathbf{x}_k) = \mu_k + \alpha_k \cos(2\pi f i + \phi_k), \tag{2.4}$$

for $i = k-T+1, k-T+2, \ldots, k$, and $\mathbf{x}_k = [\mu_k, \alpha_k, \phi_k]^t$. Note that Y_i in (2.3) is a function of only time i, for a single position k of the window and $\mathbf{x}_k$ is constant over the whole window for that position. Value of $\mathbf{x}_k$, for a single window positioned at k, can be found by finding a solution to the following equation:

$$\mathbf{x}_k = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \sum_{i=k-T+1}^{k} (y_i - F(i; \mathbf{x}_k))^2 \right\}. \tag{2.5}$$

We used MATLAB default non-linear least squares with *Trust Region Reflective* optimization algorithm [58, 59] to solve the problem in (2.5). The solution of (2.5) at every position $k = T, \ldots, N$ of the window produces parameter time-series of (2.1), $\mathbf{x}_k$, for $k = T, \ldots N$. Our study [52] about changes introduced by beetle infestation in vegetation index time-series, and findings in [22] suggest that beetle infestation affects trend component of the signal significantly and the change can be detected in the trend component alone. Therefore, we apply the change detection strategy on μ_k time-series.

We assume that we have sufficient length (at least 5 years or 5 cycles) of the stationary

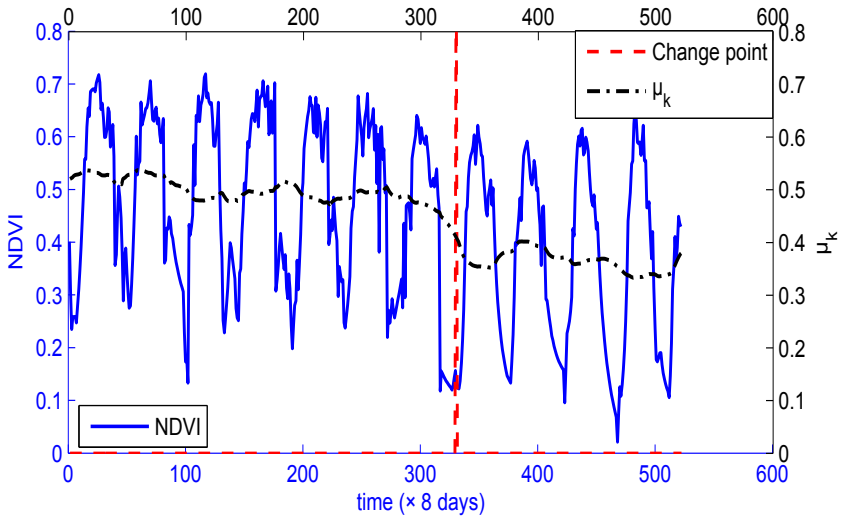

Figure 2.1: NDVI time-series containing a change plotted on left vertical and bottom horizontal axes (complete line). μ_k parameter time-series derived using moving window model-based method plotted on right vertical and top horizontal axes (dash-dot), and the detected change at time point $t = 325$ (dashed line), using window size $T = 46$, length of reference period $L = 230$, and threshold $\lambda = 4$.

time-series, during which the forest was healthy and no change occurred. Let us represent the stationary part of the vegetation index time-series by $[Y_1, \ldots, Y_L]$, where L is the length of the stationary part (5year/cycles $\times$ 46 observations per year = 230 in our case) of the vegetation index time-series. Since, μ_k is a trend component and should ideally remain constant for a noise-free vegetation index time-series, as long as the window in (2.3) and (2.5) is sliding over the stationary part, or as long as $k \leq L$. However, the vegetation index signal is not noise-free in practical cases, so we expect μ_k corresponding to the stationary vegetation index time-series to have small variations and maintain some (unknown) distribution. We assume that μ_k maintains a Gaussian distribution with mean $M_1 = \text{mean}(\mu_T, \ldots, \mu_L)$ and standard deviation $\sigma = \text{std}(\mu_T, \ldots, \mu_L)$ during the ref-

erence/stationary period and the likelihood of M_1 at time point k can be given by,

$$P_{M_1}(\mu_k) = \frac{1}{\sqrt{2\pi}\sigma} e^{\left(-\frac{(\mu_k - M_1)^2}{2\sigma^2}\right)}. \tag{2.6}$$

Our hypothesis here is that the μ_k will start deviating from the distribution of $[\mu_T, \ldots, \mu_L]$ and enter a new distribution with mean $M_2 = M_1 + \lambda\sigma$, as soon as the sliding window starts entering the non-stationary part (after infestation has occurred) of the time-series. In other words, the likelihood of M_1 will start decreasing and the likelihood of M_2, $P_{M_2}(\mu_k)$, will start rising i.e. the likelihood ratio $\frac{P_{M_1}(\mu_k)}{P_{M_2}(\mu_k)}$ will start decreasing. The log likelihood ratio can be written as,

$$\mathscr{L}_k = \ln\left(\frac{P_{M_1}(\mu_k)}{P_{M_2}(\mu_k)}\right) \tag{2.7}$$

$$\implies \mathscr{L}_k = \ln\left(\frac{\frac{1}{\sqrt{2\pi}\sigma} e^{\left(-\frac{(\mu_k - M_1)^2}{2\sigma^2}\right)}}{\frac{1}{\sqrt{2\pi}\sigma} e^{\left(-\frac{(\mu_k - M_2)^2}{2\sigma^2}\right)}}\right) \tag{2.8}$$

$$\implies \mathscr{L}_k = \frac{(\mu_k - M_2)^2 - (\mu_k - M_1)^2}{2\sigma^2}. \tag{2.9}$$

At the point of change i.e the point where the likelihood ratio starts decreasing, the slope of the log likelihood ratio will be negative. Therefore, the condition for detecting a change can be found by solving

$$\frac{d\mathscr{L}_k}{dM_2} < 0, \tag{2.10}$$

$$\implies \frac{M_2 - \mu_k}{\sigma^2} < 0, \tag{2.11}$$

$$\implies M_1 + \lambda\sigma - \mu_k < 0, \tag{2.12}$$

$$\implies \mu_k - M_1 > \lambda\sigma. \tag{2.13}$$

Algorithm 2.1 MBCD ($[Y_1, Y_2 \ldots, Y_L, \ldots, Y_k]$, T, L)

Given $Y_1, Y_2, \ldots, Y_L, \ldots, Y_k$, for some current time point $k > L$, where L is the length of the stationary period,

1. Evaluate $\mathbf{x}_t$ for all $t = T, \ldots, k$ using (2.3) to (2.5).
2. Evaluate $M = \text{mean}(\mu_T, \ldots, \mu_L)$ and $\sigma = \text{std}(\mu_T, \ldots, \mu_L)$.
3. Evaluate (2.15) and then (4.11) for the current time point, k.
4. If c given by (4.11) exists, then declare c as change point.
5. Otherwise, declare no-change event.

The change detection criterion can then be formulated to identify the change point c with in the parameter time-series as

$$c = \min \left\{ k \in \{T, \ldots, N\} : \left(\sum_{j=k-9}^{k} z_j \right) > 6 \right\}, \tag{2.14}$$

where for all $k = T, \ldots, N$,

$$z_k = \begin{cases} 1 & \text{iff } s \times (\mu_k - M_1) > \lambda \sigma \\ 0 & \text{otherwise,} \end{cases} \tag{2.15}$$

Here z_k is the initial change alarm which is raised if the deviation of the μ_k from the mean M_1 becomes greater than the threshold λ times the standard deviation σ, and indicates a potential change point. The threshold λ can be selected from the table of standard normal distribution depending upon the percentage of the tail considered as outlier. The value of λ can be tuned in order to find an acceptable trade off between false alarms, true positives and detection delay. The variable s in (2.15) can take values of either +1 or -1, depending upon the type of vegetation index under consideration. Some vegetation indices, such as RGI, change in positive direction with forest degradation [21], whereas some vegetation indices, such as NDVI, change in negative direction [21]. For the former

type of vegetation indices, $s = +1$, whereas for the latter type of vegetation indices, $s = -1$, should be used. After evaluation of z_k in (2.15), (4.11) declares the current time point k as the change point c if more than 6 initial alarms z_k, out the 10 most recent, were raised. It should be noted that k has been shown to take values from T to N in Equations (2.3) to (2.15), but they can be used in online mode as well because the change detection procedure is independent of the future observations at any time point k. Algorithm 2.1, here referred to as MBCD (Model-Based Change Detection), summarizes the main steps for using this method in online mode. Note that we used window size $T = 46$ here. Choice of optimal window size T will be discussed in detail in Section 2.4. An example of an NDVI time-series containing change, the μ_k parameter derived using this algorithm, and the detected change are shown in Figure 2.1.

2.3.2 FIR Filter-Based Method (Method 2)

As mentioned above, only μ_k parameter of the model (2.1) gets sufficient impact of non-stationarity/change and can be used alone in change detection, the change detection can be formulated in a computationally and mathematically simpler way. Since, we need only trend component of the signal, μ_k, rather than all the three parameters of (2.1), we can approximate μ_k by the response of a T-point moving average FIR filter to Y_k. Let H be the impulse response of an FIR filter, given by

$$H = [h_1, h_2, \ldots, h_T], \tag{2.16}$$

where

$$H[t] = h_t \tag{2.17}$$

and

$$h_t = \frac{1}{T}, \quad \forall t = 1, 2, \ldots, T. \tag{2.18}$$

Algorithm 2.2 MABCD ($[Y_1, Y_2 \ldots, Y_L, \ldots, Y_k]$, T, L)

Given $Y_1, Y_2, \ldots, Y_L, \ldots, Y_k$, for some current time point $k > L$, where L is the length of the stationary period, and T being the filter length,

1. Select an FIR filter according to (2.16) and (2.17).
2. Evaluate μ_k using (2.21).
3. Evaluate $M = \text{mean}(\mu_T, \ldots, \mu_L)$ and $\sigma = \text{std}(\mu_T, \ldots, \mu_L)$.
4. Evaluate (2.15) and then (4.11) for the current time point, k.
5. If c given by (4.11) exists, then declare c as change point.
6. Otherwise, declare no-change event.

The response, μ_k, of the filter given in (2.17), to the vegetation index signal $W = [Y_1, Y_2, \ldots, Y_k]$ at time k with $W[n] = Y_n$ for all $n = 1, 2, \ldots, k$, can be derived by

$$\mu_k = (W \star H)[k], \tag{2.19}$$

where

$$(W \star H)[k] = \sum_{r=-\infty}^{\infty} H[r]W[k-r+1], \tag{2.20}$$

and $\star$ is convolution operator. Since, we are using an FIR filter, we can re-write (2.20) as

$$\mu_k = \sum_{r=1}^{T} H[r]W[k-r+1]. \tag{2.21}$$

The idea here is that the response μ_k of the FIR filter in (2.16), to the vegetation index signal W, will remain consistent and maintain some distribution, as long as $k \leqslant L$, where L is the length of the known stationary part. Once k enters the non-stationary part, μ_k will start deviating from its distribution in the stationary part and will lose its consistency, which can be detected using (4.11) and (2.15). Since μ_k is approximation of the trend component of the vegetation index signal and should have no effects of the seasonal cycles

of the vegetation index signal, the value of T should be multiple of length of the cycle. We have used $T = 46$, which is the minimum value we could have because the cycle length in our data is 46. Values of T greater than 46 can also be used but it will bring delays in the detection process. Algorithm 2.2, here referred to as MABCD (Moving Average Based Change Detection), summarizes the main steps for using this method in online mode.

2.3.3 Near-real Time Disturbance Detection in MODIS Data (Method 3)

Recently, a method that deals with the problem of near-real time disturbance detection in MODIS time-series data, henceforth labeled as "Method 3" or "MOSUM-based method", was published in [32]. This method used least squares to fit a season-trend model to the reference/history part of the 16-day MODIS NDVI time-series, and derive its parameters. This model, with the derived parameters, is then used to estimate the NDVI values in the monitoring period. The difference sequence between the estimated values and the real observations is then used to derive MOSUM (MOving SUM) statistics. The change alarm is raised if the MOSUM statistic at the time-point under consideration falls outside the 95% boundary of the distribution derived from MOSUM statistics corresponding to the reference/history part of the time-series.

2.3.4 Experimental Setup I

Validation of multi-temporal change detection methods is often a hard task due to the lack of independent reference sources of potential changes [31]. After personal communication with the authorities [62–64], we came to know that the aerial surveys and the year-wise maps showing regions of infestations were not exact and could have offsets from the exact locations of the infestations. Furthermore, we could only know the period of no infestation but not when exactly the infestation occurred. The objective of this ex-

perimental setup was to generate a dataset of simulated time-series with noise from the real world dataset, in which we could introduce the changes of desired intensity at desired time-points, in order to check the sensitivity and *detection delay* of the proposed methods. The time-series were simulated following the similar procedure as used in [30–32]. However, the noise used in this experiment was derived from the real world RGI dataset.

First, the model given in (2.1) was fitted to each time-series in the real world RGI dataset and the error sequence for each time-series of the dataset was found by

$$\varepsilon_k = Y_k - (\mu_k + \alpha_k \cos(\omega k + \phi_k)). \tag{2.22}$$

Then, ε_k found for all the time-series in the dataset were combined together to form an array Q, of length q, containing all the error values of the real world RGI dataset. Secondly, the seasonal component of each cycle of the time-series was generated by an asymmetric Gaussian function [30], given by

$$g(l) \equiv g(l;a,b,\rho_1,\rho_2) = a \times \begin{cases} \exp\left[-\frac{(l-b)^2}{\rho_1}\right], & \text{if } l > b \\ \exp\left[-\frac{(b-l)^2}{\rho_2}\right], & \text{if } l < b, \end{cases} \tag{2.23}$$

where a, b are the amplitude and the position of the maximum or minimum with respect to the time l, whereas ρ_1 and ρ_2 determine the width of the left and right hand sides, respectively. Note that the values of a, b, ρ_1 and ρ_2 can be selected according to the type of dataset one wants to simulate. For example, we are using MODIS 8-days composite data which means cycle length of 46 time-points, we used $a = 0.7$, $b = 23 + \lfloor l/46 \rfloor \times 46$ and $\rho_1 = \rho_2 = 100$ in (2.23), in order to mimic the seasonal component of cycle length of 46 time-points and make the first cycle centered on $l = 23$, second on $l = 23 + 46 = 69$ and so on. Since, our methods do not depend upon the absolute values of trend and seasonal cycle amplitude, and look for changes in a parameter time-series, by comparing

the current parameter value with the distribution of the parameter values during the history (*normal*) period, the value of a can be selected arbitrarily within a realistic range of the vegetation index being simulated. The simulated time-series was then generated by

$$S(l) = g(l) + \psi(l) + Q[\vartheta_l] \tag{2.24}$$

where $Q[\cdot]$ is the noise/error value at position $(\cdot)$ of the noise/error array Q and ϑ_l is an integer random number drawn from a uniform distribution at time l, such that $1 \leq \vartheta_l \leq q$. The second term in the right hand side of (2.24), $\psi(l)$, is the simulated change introduced in the time-series, and can be given by

$$\psi(l) = \begin{cases} \left[(1)^{\lfloor l/\rho \rfloor} - (0)^{\lfloor l/\rho \rfloor}\right] \times \varsigma \times (l-\rho) & \text{if } l \leq \xi \\ \psi(\xi) & \text{if } l > \xi, \end{cases} \tag{2.25}$$

where ς, ρ, and ξ are the slope, start point and end point of the introduced change, respectively. The values of these parameters can be set by the user, depending upon the desired intensity and length of the change. For generating time-series without any changes (no-change examples), the $\psi(l)$ term in (2.24) was made zero. Note that it is a challenging task to simulate a remotely sensed time-series with combined information on vegetation phenology, inter-annual variability, disturbance events and signal contamination [31]. Therefore, testing these methods on real world remotely sensed data is necessary.

2.3.5 Experimental Setup II

The objective of this experimental setup was to test the performance and robustness of both the methods against different magnitudes of the signal noise. The time-series were generated following the procedure explained in Section 2.3.4, except the noise used in (2.24). The modified equation (2.24), used for generating change and no-change examples

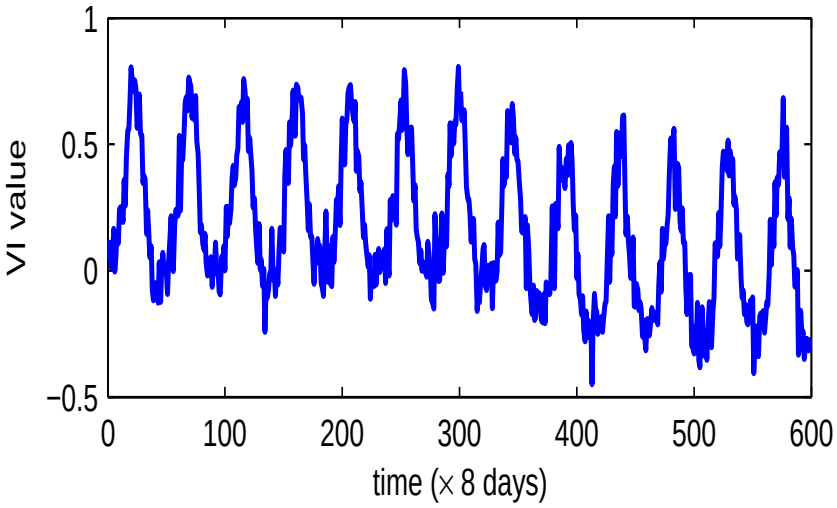

Figure 2.2: Simulated time-series containing gradual change at $l = 301$, generated by experimental setup II, using $\sigma_{noise} = 0.08$, $a = 0.6$, $b = 23 + \lfloor l/46 \rfloor \times 46$, $\rho_1 = \rho_2 = 100$, $\varsigma = 0.002$, $\rho = 301$, and $\xi = 400$.

in this simulation experiment, is given by

$$S(l) = g(l) + \psi(l) + v_l \tag{2.26}$$

where v_l is the noise value drawn randomly from a normal distribution, $N(0, \sigma_{noise})$. All the other equations given in Section 2.3.4 were used unchanged. An example of a simulated time-series generated by experimental setup II with σ_{noise}=0.08, containing gradual change of slope 0.002 at $l = 301$, is shown in Figure 2.2.

2.3.6 Beetle Infestation Data

We prepared the datasets consisting of 355 pixels which were infested after 2005 and 355 which were never infested until the end of 2011. These pixels were taken from the Rocky Mountain region (Colorado), Inter-mountain Region (Utah) in the United States and the northern parts of British Columbia, Canada. First the government survey maps from forest

Table 2.1: Different vegetation indices (VI) used in this study, and the associated formulas.

VI	NDVI	GNDVI	WDRVI	NDII6	NDII7	SR	ISR6	ISR7	RGI	NDWI
Formula	$\frac{B_2-B_1}{B_1+B_2}$	$\frac{B_2-B_4}{B_4+B_2}$	$\frac{(\alpha+1)NDVI+(\alpha-1)}{(\alpha-1)NDVI+(\alpha+1)}$	$\frac{B_2-B_6}{B_4+B_6}$	$\frac{B_2-B_7}{B_4+B_7}$	$\frac{B_1}{B_2}$	$\frac{B_6}{B_2}$	$\frac{B_7}{B_2}$	$\frac{B_1}{B_4}$	$\frac{B_2-B_5}{B_2+B_5}$

health services [62–64] were used as references to identify the regions on the Google Earth, which were infested after 2005 and those with no infestations ever recorded. Then the geographical coordinates of the desired regions were recorded after visual verification on Google Earth. These coordinates were then used in an online tool, MODLAND Tile Calculator [65], to find their respective MODIS horizontal and vertical tile numbers in the MODIS sinusoidal grid. One MODIS image of 500m spatial resolution was downloaded for each of the relevant tiles. These MODIS images were used in ENVI (version 5) in order to find out the pixel coordinates of the selected regions within the respective MODIS tile, using their recorded geographical coordinates. Then, the time-series datasets of the selected pixels were prepared from the MODIS product MCD43A4.005 using different vegetation indices, as shown in the Table 2.1.

2.4 Results and Discussion

2.4.1 Results for Simulation Experiment 1

In simulation experiment 1, 550 change and 550 no-change examples were generated following the process explained in Section 2.3.4. The seasonal cycles were generated using asymmetric Gaussian function, the change was introduced by adding a ramp of slope 0.001 ($\varsigma = 0.001$ in (2.25)) to the signals at various positions, and the noise introduced into the signal was drawn randomly from the noise distribution of the real RGI dataset. Both the methods were analyzed and compared on the simulated dataset, using different thresholds. The detailed comparison has been plotted in Figure 2.3. Table 2.2 shows a comparison between the accuracies and the corresponding mean detection delays

Table 2.2: Experimental Setup 1: Performance comparison of Method 1 and Method 2 on simulated data with noise from RGI dataset, at acceptable accuracy and mean delay in detections. Method 1= Model-based method, Method 2= FIR-based method, TP = True Positive, TN = True Negative, Acc. = Overall Accuracy, MD = Mean Delay in detections, δ = mean size of change detected and λ = threshold value.

Method	TP (%)	TN (%)	Acc. (%)	MD	δ	λ
1	93.09	86.90	90	54.4	0.054	3.76
2	92.90	87.09	90	54.4	0.054	3.82

achieved by both the methods. For the same overall accuracies, both the methods have the same mean detection delays and the corresponding sizes of the detected changes. Table 2.3 compares the maximum accuracies achieved by both the methods, their corresponding mean delays at these accuracies and the change sizes detected by them. Both the methods show almost equal performance indices at the same maximum accuracies. Both of them utilize full temporal resolution of the signal, which is why the changes have been detected very shortly after time of occurrence. Note that the mean delay (MD) values of 54.4 and 66 should not be taken as absolute delays here because the slope of the introduced changes is very small (1×10^{-3}). A delay of 54 points means, the change is detected when its magnitude is $1 \times 10^{-3} \times 54 = 0.054$. It is clear from the Tables (2.2 and 2.3) that both the methods can detect changes with high accuracy, TP and TN. Although, both the methods achieve similar performance on simulated data with noise from the real dataset, comparison of these methods on real world dataset is also required. The real world datasets involve much more complex phenomena, and changes are not defined as clearly as in the simulated dataset. The results here encourage us to consider both the methods for testing with the real world datasets (see Section 2.4.3).

2.4.2 Results for Simulation Experiment 2

In simulation experiment 2, each time-series was generated in exactly the same way as mentioned in Section 2.3.5. We generated 20 datasets for 20 different values of noise

Table 2.3: Experimental Setup 1: Performance comparison of Method 1 and Method 2 on simulated data with noise from RGI dataset, at maximum accuracy. Method 1= Model-based method, Method 2= FIR-based method, TP = True Positive, TN = True Negative, Acc. = Overall Accuracy, MD = Mean Delay in detections, δ = mean size of change detected and λ = threshold value.

Method	TP (%)	TN (%)	Acc. (%)	MD	δ	λ
1	99.45	98.72	99.1	66.1	0.066	5.52
2	99.63	98.54	99.1	65.35	0.065	5.56

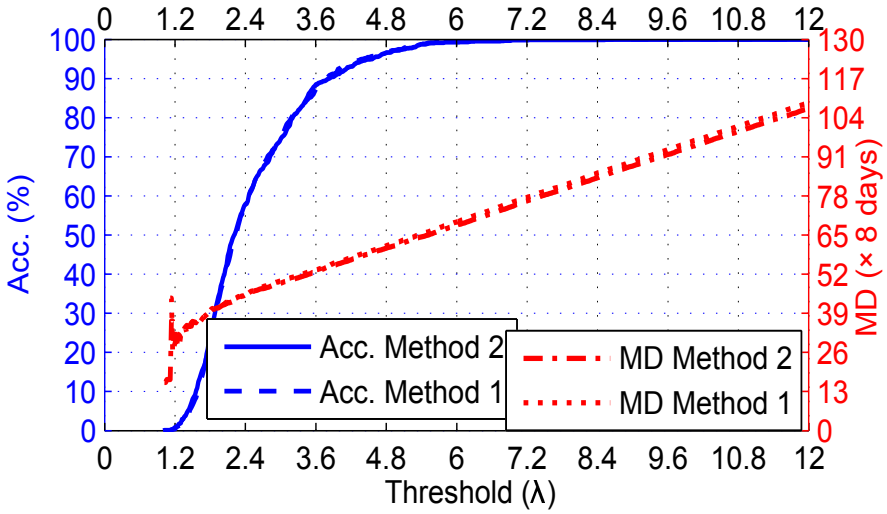

Figure 2.3: Comparison between performances of model-based and moving average change detection methods with different thresholds, on simulated data with noise from real RGI dataset. Threshold vs. Accuracy (plotted on left vertical and bottom horizontal axes). Threshold vs. mean detection delay (plotted on right vertical and top horizontal axes).

standard deviation σ_{noise}, starting from $\sigma_{noise} = 0.01$ to as high as $\sigma_{noise} = 0.2$. Each dataset contained 300 change and 300 no-change examples, with known positions of the introduced gradual changes and their slopes ($\varsigma = 0.001$ in (2.25)). This analysis was meant to test the sensitivity and robustness of both the methods, against different magnitudes of noise. A detail comparison between performances of both the methods, for simulated datasets of different noise standard deviations (σ_{noise}), is plotted in Figure 2.4. The accuracy of the model-based method is higher than the FIR-based method for noise with $\sigma_{noise} \leq 0.7$. Both the accuracy curves overlap each other and do not drop much for higher values of σ_{noise}. Both the methods achieve high accuracies (above 90%) even when the introduced noise is as big as of standard deviation of 0.2. However, the mean detection delay and the size of the detected change have increased with the increase in σ_{noise}. This is because in case of high values of noise, the change needs to be significant in order to get detected. Performances of both the methods are the same in most of the cases. Both can detect a change of magnitude 0.122 (slope $\times$ delay=0.001 $\times$122) with nearly 95% accuracy, when the noise in the original time-series is as big as of the standard deviation of 0.2.

2.4.3 Results for Real Datasets

We applied both the methods on different vegetation index datasets of the real 355 change and 355 no-change examples. Table 2.4 shows the comparison between the accuracies, and the corresponding mean detection delays, achieved by both the change detection methods, for different vegetation indices. The model-based method detected the changes much faster than the FIR-based method for similar or more accuracy. This difference in the performance can be attributed to the fact that the moving average FIR filter is more sensitive to noise of the original vegetation index signal and hence, detects the change only when it becomes significant enough. However, in case of model-based method, the effect of noise of the original signal is not translated fully into the μ_k parameter time-series

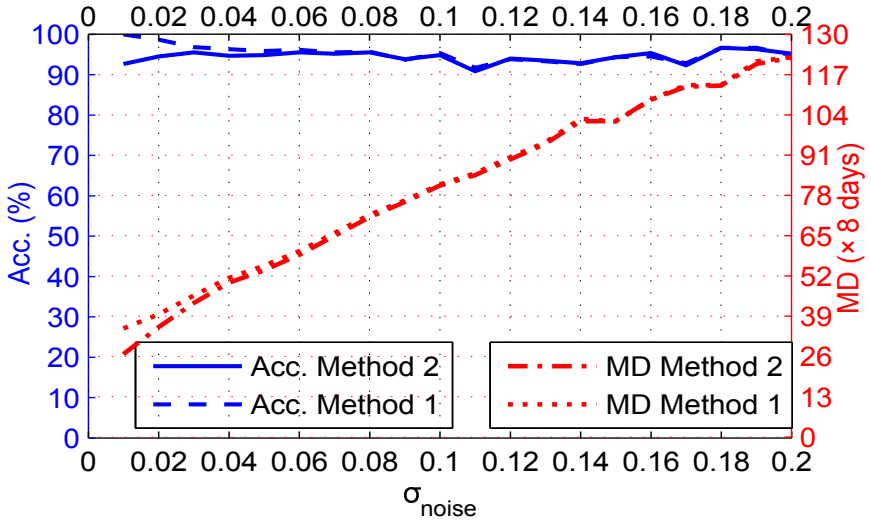

Figure 2.4: Comparison between performances of model-based and moving average change detection methods, on simulated data with different magnitudes of noise from normal distributions. Noise standard deviation (σ_{noise}) vs. Accuracy (plotted on left vertical and bottom horizontal axes). Noise standard deviation (σ_{noise}) vs. Mean Detection Delay (plotted on right vertical and top horizontal axes).

because some of it is absorbed by the α_k parameter as well.

A comparison between the performance indices of the model-based method on different vegetation indices can be observed in the left half of the Table 2.4. Comparison between the first four and rest of the rows in the left half of the table shows that the *ratio-indices* i.e. RGI, SR, ISR6 and ISR7, detect the changes much faster than the *difference-indices* i.e. NDVI, NDWI, WDRVI, GNDVI, NDII6 and NDII7, with the same accuracies. Figure 2.5 shows a detailed comparison between the accuracies and the corresponding mean detection delays for the four ratio-indices at different threshold values. At start, the accuracy curves of ISR6 and ISR7 are slightly higher than the curves for SR and RGI. But, this difference is not big and becomes very small at the threshold value $\lambda = 4$. However, the corresponding curves for mean detection delays have significant difference at every point. The mean detection delay curve for RGI is significantly lower than the curves for the rest of the indices.

Table 2.4: Comparison between performances of Model-based method (method 1) and FIR-based method (Method 2), on real datasets of different indices, in terms of corresponding mean delays in detections, at acceptable accuracies. TP = True Positive, TN = True Negative, Acc. = Overall Accuracy, MD = Mean Delay in detections, and λ = threshold value.

Vegetation Index	Method 1					Method 2				
	TP (%)	TN (%)	Acc. (%)	MD	λ	TP (%)	TN (%)	Acc. (%)	MD	λ
RGI	92.7	83.5	88.1	67.0	3.58	91.3	85.0	88.2	86.2	2.38
SR	98.9	77.5	88.1	73.0	2.94	96.0	80.0	88.0	88.0	1.92
ISR6	92.5	83.5	88.0	71.0	2.40	82.6	93.4	88.0	102.7	3.14
ISR7	94.7	81.5	88.0	72.4	2.20	85.4	90.5	87.9	102.5	2.66
NDVI	97.8	78.5	88.2	82.9	3.50	95.8	80.0	87.9	86.5	1.94
NDWI	83.1	71.0	77.1	94.1	3.12	64.51	63.5	64.0	103.1	2.00
WDRVI	96.6	79.5	88.1	85.5	4.38	97.7	77.5	87.6	90.0	2.18
GNDVI	94.7	81.5	88.1	106.5	3.62	94.64	81.0	87.8	107.8	2.16
NDII6	89.6	86.5	88.1	86.0	3.14	82.0	93.5	87.7	110.8	3.36
NDII7	93.8	82.5	88.2	85.6	2.46	85.1	90.5	87.8	103.7	2.82

The best choice of vegetation index to be used for change detection depends upon the user's preference guided by the trade-off between the detection accuracy and the detection delay. If real time detection is not crucial, then ISR6 can be used to achieve the best detection accuracy (see Figure 2.5). However, the RGI provides the best trade-off between the accuracy and the detection delay (see Figure 2.5). The detection delay can be reduced considerably at the expense of a very slight decrease in the accuracy by using RGI.

2.4.4 Results for Different Window Sizes Using Model-Based Method

The size, T, of the sliding window used in model-based change detection method should be selected carefully. Too small, a window size, will give low accuracy and higher delays in the detection. Too big window sizes will achieve better accuracies but delayed detections. This has been shown in the Figure 2.6. The accuracy of detection increases rapidly with the increase in the window sizes until it reaches its maximum at the size of 40. Af-

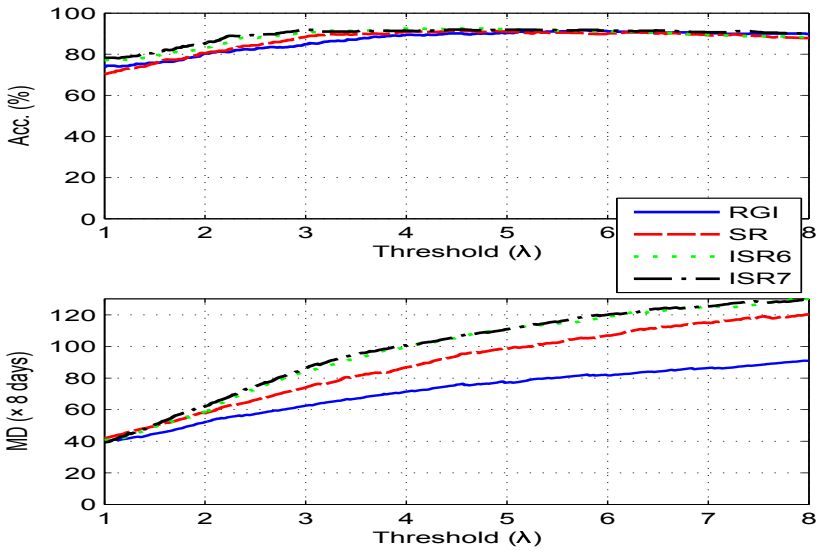

Figure 2.5: Performance of model-based change detection method for different vegetation index datasets, at different thresholds. Threshold vs. Accuracy (Top) and Threshold vs. mean detection delay (Bottom).

ter 40 the accuracy remains almost constant with slight variations, and start decreasing slowly after the size becomes greater than 80. The mean detection delay falls rapidly with increase in the window size, until it reaches its lowest value at 46 and starts increasing again slowly. The low accuracies and higher detection delays for smaller sized windows are because the model-fitting is not optimal. This also makes the method more prone to the noise in the original time-series, which makes it difficult to detect the changes quickly and accurately. The window sizes between 46 and 80 achieve high accuracies. However, the method becomes less sensitive to the changes and takes longer to detect them, with increasing window sizes. Figure 2.6 shows that the window size equal to seasonal cycle length (46 in our case, shown by the arrows in the figure) provides optimal trade-off between the accuracy and the mean detection delay. This can be attributed to the fact that a window of size 46 covers exactly one cycle of the time-series, caters for all the seasonal

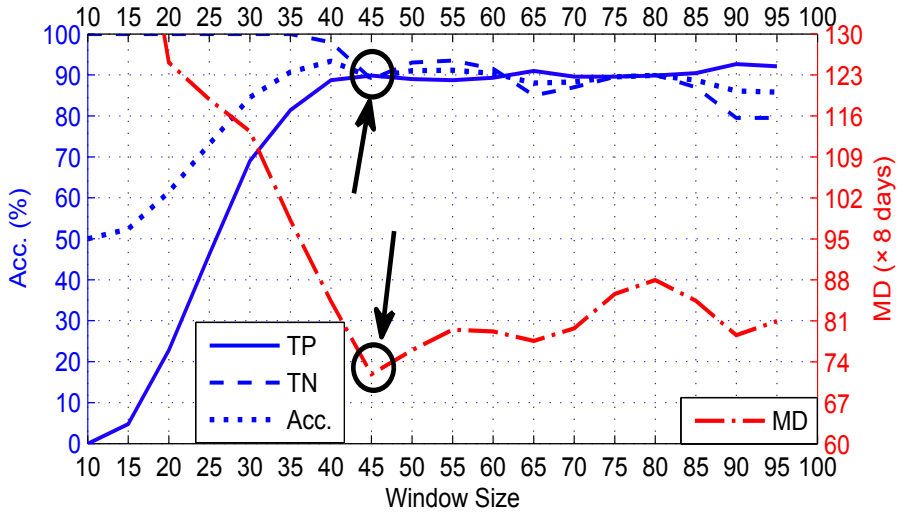

Figure 2.6: Performance of model-based change detection method with different window sizes. Window-sizes vs. Accuracy, TP and TN (plotted on left vertical and bottom horizontal axes) and Window-sizes vs. mean detection delay (plotted on right vertical and top horizontal axes). Arrows highlighting the performance with optimal window size.

variations, and makes the model-fitting optimal.

2.4.5 Cross Validation Results of the Model-Based Method

The threshold used in model-based method can be tuned according to the users requirement whether fast detection is crucial or only high accuracy is required. To ensure the robustness of the selected threshold for test data, we cross-validated the threshold selection on RGI dataset. In every experiment, we selected the thresholds corresponding to the points where TP and TN were the closest and TP was greater than the TN.

Table 2.5 shows the results of cross-validations using 90%, 50% and 25% of the dataset as training sets, and 10 %, 50% and 75% of the dataset as the respective test sets. The table has 3 major parts. Right half of the first two rows in each part provides a comparison between the mean performance indices (TP, TN, Acc. and MD) of the training and test sets after 10 runs, whereas the right half provides a comparison between the corresponding

Table 2.5: Results of three cross-validations of Method 1, on the RGI dataset. In the three experiments, 10%, 50% and 75% of the dataset were used as test sets, and 90%, 50% and 25% of the dataset were used as training sets, respectively. TP = True Positive, TN = True Negative, Acc. = Overall Accuracy, MD = Mean Detection Delay, and λ = threshold value.

S.No.	Data	Mean (10 runs)				Standard Deviation (10 runs)				% of the Dataset
		TP (%)	TN (%)	Acc. (%)	MD	TP (%)	TN (%)	Acc. (%)	MD	
1	Training	89.8	89.3	89.5	73.0	0.14	0.14	0.14	0.40	90
	Test	89.5	89.5	89.5	72.8	1.80	2.50	1.25	0.94	10
	λ	4.22				0.04				NA
2	Training	89.7	89.1	89.4	73.1	0.25	0.29	0.26	1.10	50
	Test	89.5	89.4	89.4	73.6	0.90	1.50	0.41	1.04	50
	λ	4.22				0.13				NA
3	Training	90.3	89.3	89.8	72.2	0.74	0.73	0.26	0.92	25
	Test	89.8	88.6	89.2	72.6	0.75	1.34	0.49	1.25	75
	λ	4.18				0.13				NA

standard deviations. The third row in each part of the table gives the mean and standard deviation of the thresholds selected in the 10 runs of the respective cross-validation experiment. The last column of the table shows the percentages of the original dataset which were used as training or test sets in each of the cross-validation experiments. The results in the table show that the corresponding mean performance indices are close to each other for training and test sets. The standard deviations of the performance indices of the training set are smaller than those of the test set, which is expected in any cross-validation experiment. The standard deviations of the performance indices in all the three cross-validation experiments are not very big. Furthermore, the corresponding mean values of the performance indices across all the three cross-validation experiments are very close. These facts indicate that the selected thresholds for RGI dataset are robust for test data.

2.4.6 Comparison Between Method 1 and Method 3

We showed above that the method 1 (model-based method) performs better than the method 2 (FIR filter-based method) for near-real time detection of changes in MODIS

time series data. For further performance evaluation of method 1, we compared it with a recently published method for near-real time disturbance detection in MODIS 16-day data (method 3) [32], briefly explained in Section 2.3.3, on both simulated as well as real world datasets. The results are explained as follows.

First we compared the results of method 1 and method 3 on the dataset generated by experimental setup II, which is similar to the one also used in [32]. The comparisons between the accuracies and the mean detection delays achieved by both the methods have been shown in Figure 2.7. It can be noticed that method 1 is more robust to noise than method 3, and achieves better accuracies with lower mean detection delays at almost all the noise levels.

The performances of both the methods were also compared on different real world vegetation index datasets. The comparison of best trade-offs between accuracies and detection delays, achieved by both the methods for different vegetation index datasets, is shown in Table 2.6. By comparing the columns in the left half of the table with the corresponding columns in the right half of the table, it can be noticed that method 1 achieved higher TP and TN with lower mean detection delays than method 3, for all the vegetation index datasets. Please note that due to the lack of information about the actual change time-points, the mean detection delays were calculated from a fixed time-point (end of year 2005, L=230), till which, we knew that there had been no infestation reported. Therefore, the high values of mean detection delays should not create any confusion as they are meant solely for comparison purpose and should not be taken in absolute sense.

These results revealed that detecting changes in the parameter time-series, derived by fitting a triply modulated cosine model over a moving window, performs better than detecting changes in the noise time-series formed by difference between the observations and predicted vegetation index values. This can be attributed to the fact that the least square fitting, of the model over a sliding window, reduces the effect of noise and the

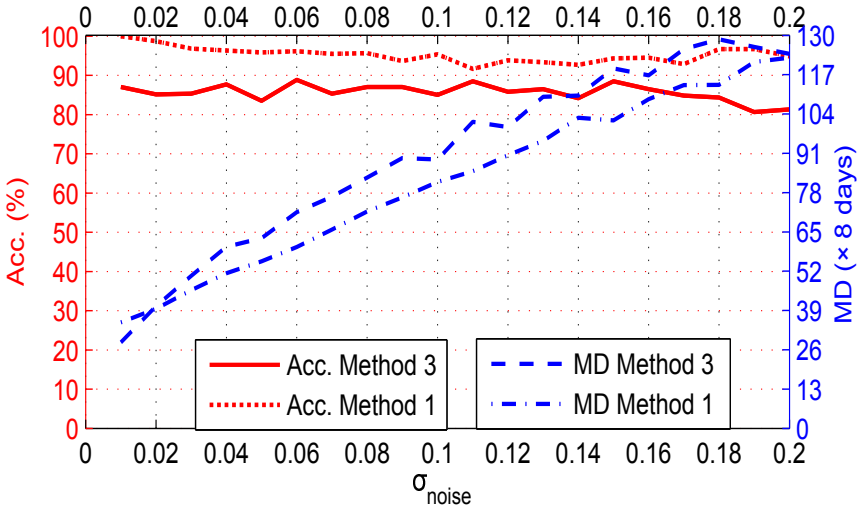

Figure 2.7: Comparison between performances of model-based (Method 1) and MOSUM-based (Method 3) change detection methods, on simulated data with different magnitudes of noise from normal distributions. Noise standard deviation (σ_{noise}) vs. Accuracy (plotted on left vertical and bottom horizontal axes) and Noise standard deviation (σ_{noise}) vs. mean detection delay (plotted on right vertical and top horizontal axes).

derived trend parameter provides clear separation/shift between the stationarity and non-stationarities. On the other hand, the noise sequence or the difference sequence between the model and the real observations, has variations of wider span, and as a result the range of noise values from the stationary part can overlap the range of noise values from the non-stationary part. This makes it harder to detect the changes early.

2.5 Conclusions

We proposed two methods for near-real time detection of beetle infestation in pine forests using MODIS 8-days 500m data. Although some methods for change detection do exist in literature [10, 21, 22, 27, 29, 49], there are currently no methods, specifically for near-real time detection of beetle infestation, which utilize full temporal resolution of MODIS 8-days 500m data. The proposed methods were tested on simulated data, data with different

Table 2.6: Comparison between performances of Model-based method (Method 1) and MOSUM-based method (Method 3), on real datasets of different indices, in terms of corresponding mean delays in detections, at acceptable accuracies. TP = True Positive, TN = True Negative, Acc. = Overall Accuracy, MD = Mean Delay in detections, and λ = threshold value.

Vegetation Index	Method 1					Method 3				
	TP (%)	TN (%)	Acc. (%)	MD	λ	TP (%)	TN (%)	Acc. (%)	MD	λ
RGI	92.7	83.5	88.1	67.0	3.58	86.2	64.5	75.4	96.0	4.50
SR	98.9	77.5	88.1	73.0	2.94	91.7	64.5	78.6	103.0	3.60
ISR6	92.5	83.5	88.0	71.0	2.40	75.8	73.0	74.4	88.5	2.20
ISR7	94.7	81.5	88.0	72.4	2.20	77.2	76.5	76.8	106.5	2.34
NDVI	97.8	78.5	88.2	82.9	3.50	92.4	63.5	77.9	110.4	4.00
NDII6	89.6	86.5	88.1	86.0	3.14	77.5	68.5	73.0	83.1	2.14
NDII7	93.8	82.5	88.2	85.6	2.46	75.8	75.6	75.6	99.8	2.14

noise levels, and real world dataset. The results suggest that both the methods achieve high detection accuracy with high true positive and true negative rates. However, the model-based method detects changes faster and suits the detection of beetle infestation in near-real time. So, if near-real time detection is not required, these findings suggest using finite impulse response filter-based (moving average) method because it is easier to implement, and computationally simpler. The results also show that the model-based method detects changes quicker with the *ratio-indices* than with the *difference- indices*. Among the ratio-indices, RGI provides the best trade-off between detection accuracy and detection delay, for the study region under consideration. A window size equal to the length of the cycle of the vegetation index signal gives optimal performance in the model-based method. The cross-validation results show that the threshold selection is robust and performs accurately on test data. The comparison with an existing near-real time disturbance detection technique showed that detecting change in the parameter time-series is more accurate and faster than detecting it in the noise time-series derived from the difference between the modeled values and real observations.

These findings suggest that MODIS 8-days composite 500m data suits the detection of

beetle infestation. The utilization of full temporal resolution of this MODIS product allows us to detect beetle infestation in near-real time, which is an advantage of such high temporal resolution data. Also, detecting changes in trend parameter time-series suits beetle infestation detection problem. Since, the method is purely of statistical nature, it may detect other types of changes, which affect the statistics of the parameter time-series, as well. Nevertheless, it can serve as an initial alarm system to flag the areas with high probability of having unsual activity (change), which can then be inspected more closely by the concerned authorities. Although, the results achieved for North American beetle infestation are impressive and encouraging, we recommend testing these methods and MODIS 8-days 500m data in other regions as well.

Bibliography

[1] R. Watson, I. Noble, B. Bolin, N. Ravindranath, D. Verardo, and D. Dokken, "Land Use, Land-Use Change and Forestry," *Cambridge University Press, UK*, p. 375, 2000.

[2] B. Bentz and D. Endreson, "Evaluating satellite imagery for estimating mountain pine beetle-caused lodgepole pine mortality: Current status," *Mountain Pine Beetle Symposium: Challenges and Solutions*, p. 298, 2003.

[3] S. Bontemps, P. Bogaert, N. Titeux, and P. Defourny, "An object-based change detection method accounting for temporal dependences in time series with medium to coarse spatial resolution," *Remote Sensing of Environment*, vol. 112, pp. 3181–3191, Jun 2008.

[4] J. S. Borak, E. F. Lambin, and A. H. Strahler, "The use of temporal metrics for land cover change detection at coarse spatial scales," *International Journal of Remote Sensing*, vol. 21, pp. 1415–1432, Jan 2000.

[5] N. Bories, J. Samalens, D. Guyon, N. Breda, and J. Wigneron, "Monitoring pine defoliation due to the processionary moth at regional scale from modis time series," in *International Geoscience and Remote Sensing Symposium (IGARSS), IEEE*, pp. 3383–3386, 2012.

[6] N. C. Coops, M. Johnson, M. A. Wulder, and J. C. White, "Assessment of quickbird

high spatial resolution imagery to detect red attack damage due to mountain pine beetle infestation," *Remote Sensing of Environment*, vol. 103, pp. 67–80, Jul 2006.

[7] N. C. Coops, S. N. Gillanders, M. A. Wulder, S. E. Gergel, T. Nelson, and N. R. Goodwin, "Assessing changes in forest fragmentation following infestation using time series landsat imagery," *Forest Ecology and Management*, vol. 259, pp. 2355–2365, May 2010.

[8] P. Coppin, I. Jonckheere, K. Nackaerts, and B. Muys, "Digital change detection methods in ecosystem monitoring: A review," *Int. J. Remote Sens.*, vol. 25, pp. 1565–1596, May 2004.

[9] K. M. de Beurs and G. M. Henebry, "A statistical framework for the analysis of long image time series," *International Journal of Remote Sensing*, vol. 26, pp. 1551–1573, Apr 2005.

[10] K. M. de Beurs and P. Townsend, "Estimating the effect of gypsy moth defoliation using modis," *Remote Sensing of Environment*, vol. 112, pp. 3983–3990, Oct 2008.

[11] S. Franklin, M. Wulder, R. Skakun, and A. Carroll, "Mountain pine beetle red-attack forest damage classification using stratified landsat tm data in british columbia, canada," *Photogrammetric Engineering & Remote Sensing*, vol. 69, pp. 283–288, Mar 2003.

[12] R. H. Fraser and R. Latifovic, "Mapping insect-induced tree defoliation and mortality using coarse spatial resolution satellite imagery," *International Journal of Remote Sensing*, vol. 26, pp. 193–200, Jan 2005.

[13] R. H. Fraser, T. A. Abuelgasim, and R. Latifovic, "A method for detecting large-scale forest cover change using coarse spatial resolution imagery," *Remote Sens. Environ.*, vol. 95, pp. 414–427, Apr 2005.

[14] R. E. Kennedy, W. B. Cohen, and T. A. Schroeder, "Trajectory-based change detec-

tion for automated characterization of forest disturbance dynamics," *Remote Sensing of Environment*, vol. 110, pp. 370–386, Oct 2007.

[15] W. Kleynhans, J. C. Olivier, K. J. Wessels, F. van den Bergh, B. P. Salmon, and K. C. Steenkamp, "Improving land cover class separation using an extended kalman filter on modis ndvi time-series data," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, pp. 381–385, Apr 2010.

[16] W. Kleynhans, J. C. Olivier, K. J. Wessels, B. P. Salmon, F. van den Bergh, and K. Steenkamp, "Detecting land cover change using an extended kalman filter on modis ndvi time-series data," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, pp. 507–511, May 2011.

[17] E. Lambin, "Change detection at multiple temporal scales: Seasonal and annual variations in landscape variables," *Photogrammetric Engineering and Remote Sensing*, vol. 62, pp. 931 – 928, August 1996.

[18] S. Lhermitte, J. Verbesselt, W. Verstraeten, and P. Coppin, "A comparison of time series similarity measures for classification and change detection of ecosystem dynamics," *Remote Sensing of Environment*, vol. 115, pp. 3129–3152, Dec 2011.

[19] R. S. Lunetta, J. F. Knight, J. Ediriwickrema, J. G. Lyon, and L. D. Worthy, "Land-cover change detection using multi-temporal modis ndvi data," *Remote Sens. Environ.*, vol. 105, pp. 142–154, Nov 2006.

[20] A. J. Meddens, J. A. Hicke, and L. A. Vierling, "Evaluating the potential of multispectral imagery to map multiple stages of tree mortality," *Remote Sensing of Environment*, vol. 115, pp. 1632–1642, Jul 2011.

[21] A. J. Meddens, J. A. Hicke, L. A. Vierling, and A. T. Hudak, "Evaluating methods to detect bark beetle-caused tree mortality using single-date and multi-date landsat imagery," *Remote Sensing of Environment*, vol. 132, pp. 49–58, May 2013.

[22] G. W. Meigs, R. E. Kennedy, and W. B. Cohen, "A landsat time series approach to characterize bark beetle and defoliator impacts on tree mortality and surface fuels in conifer forests," *Remote Sensing of Environment*, vol. 115, pp. 3707–3718, Dec 2011.

[23] A. A. Millward, J. M. Piwowar, and P. J. Howarth, "Time-series analysis of medium-resolution, multisensor satellite data for identifying landscape change," *Photogrammetric Engineering & Remote Sensing*, vol. 72, pp. 653–663, Jun 2006.

[24] K. P. Price and M. E. Jakubauskas, "Spectral retrogression and insect damage in lodgepole pine successional forests," *Int. J. Remote Sens.*, vol. 19, no. 8, pp. 1627–1632, 1998.

[25] B. Salmon, J. Olivier, W. Kleynhans, K. Wessels, F. van den Bergh, and K. Steenkamp, "The use of a multilayer perceptron for detecting new human settlements from a time series of modis images," *International Journal of Applied Earth Observation and Geoinformation*, vol. 13, pp. 873–883, Dec 2011.

[26] B. P. Salmon, J. C. Olivier, K. J. Wessels, W. Kleynhans, F. van den Bergh, and K. C. Steenkamp, "Unsupervised land cover change detection: Meaningful sequential time series analysis," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, pp. 327–335, Jun 2011.

[27] J. P. Spruce, S. Sader, R. E. Ryan, J. Smoot, P. Kuper, K. Ross, D. Prados, J. Russell, G. Gasser, and R. McKellip, "Assessment of modis ndvi time series data products for detecting forest defoliation by gypsy moth outbreaks," *Remote Sensing of Environment*, vol. 115, pp. 427–437, Feb 2011.

[28] J. Verbesselt, P. Jonsson, S. Lhermitte, J. van Aardt, and P. Coppin, "Evaluating satellite and climate data-derived indices as fire risk indicators in savanna ecosys-

tems," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, pp. 1622–1632, Jun 2006.

[29] J. Verbesselt, A. Robinson, C. Stone, and D. Culvenor, "Forecasting tree mortality using change metrics derived from modis satellite data," *Forest Ecology and Management*, vol. 258, pp. 1166–1173, Sep 2009.

[30] J. Verbesselt, R. Hyndman, A. Zeileis, and D. Culvenor, "Phenological change detection while accounting for abrupt and gradual trends in satellite image time series," *Remote Sensing of Environment*, vol. 114, pp. 2970–2980, Dec 2010.

[31] J. Verbesselt, R. Hyndman, G. Newnham, and D. Culvenor, "Detecting trend and seasonal changes in satellite image time series," *Remote Sensing of Environment*, vol. 114, pp. 106–115, Jan 2010.

[32] J. Verbesselt, A. Zeileis, and M. Herold, "Near real-time disturbance detection using satellite image time series," *Remote Sensing of Environment*, vol. 123, pp. 98–108, Aug 2012.

[33] J. White, M. Wulder, D. Brooks, R. Reich, and R. Wheate, "Detection of red attack stage mountain pine beetle infestation with high spatial resolution satellite imagery," *Remote Sensing of Environment*, vol. 96, pp. 340–351, Jun 2005.

[34] J. White, W. M.A., and D. Grills, "Detecting and mapping mountain pine beetle red-attack damage with spot-5 10-m multispectral imagery," *BC Journal of Ecosystems and Management*, vol. 7, no. 2, pp. 105 – 118, 2006.

[35] J. C. White, N. C. Coops, T. Hilker, M. A. Wulder, and A. L. Carroll, "Detecting mountain pine beetle red attack damage with eo-1 hyperion moisture indices," *International Journal of Remote Sensing*, vol. 28, pp. 2111–2121, May 2007.

[36] M. Wulder, J. White, B. Bentz, M. Alvarez, and N. Coops, "Estimating the proba-

bility of mountain pine beetle red-attack damage," *Remote Sensing of Environment*, vol. 101, pp. 150–166, Mar 2006.

[37] M. A. Wulder, C. C. Dymond, J. C. White, D. G. Leckie, and A. L. Carroll, "Surveying mountain pine beetle damage of forests: A review of remote sensing opportunities," *For. Ecol. Manag.*, vol. 221, no. 1-3, pp. 27–41, 2006.

[38] X. Zhan, R. Sohlberg, J. Townshend, C. DiMiceli, M. Carroll, J. Eastman, M. Hansen, and R. DeFries, "Detection of land cover changes using modis 250 m data," *Remote Sensing of Environment*, vol. 83, pp. 336–350, Nov 2002.

[39] T. Cheng, B. Rivard, G. Sánchez-Azofeifa, J. Feng, and M. Calvo-Polanco, "Continuous wavelet analysis for the detection of green attack damage due to mountain pine beetle infestation," *Remote Sensing of Environment*, vol. 114, pp. 899–910, Apr 2010.

[40] S. Chien, B. Cichy, A. Davies, D. Tran, G. Rabideau, R. Castano, R. Sherwood, D. Mandl, S. Frye, S. Shulman, J. Jones, and S. Grosvenor, "An autonomous earth-observing sensorweb," *IEEE Intelligent Systems*, vol. 20, pp. 16–24, May 2005.

[41] N. C. Coops, M. A. Wulder, and J. C. White, "Integrating remotely sensed and ancillary data sources to characterize a mountain pine beetle infestation," *Remote Sensing of Environment*, vol. 105, pp. 83–97, Nov 2006.

[42] N. R. Goodwin, S. Magnussen, N. C. Coops, and M. A. Wulder, "Curve fitting of time-series landsat imagery for characterizing a mountain pine beetle infestation," *International Journal of Remote Sensing*, vol. 31, pp. 3263–3271, Jun 2010.

[43] N. R. Goodwin, N. C. Coops, M. A. Wulder, S. Gillanders, T. A. Schroeder, and T. Nelson, "Estimation of insect infestation dynamics using a temporal sequence of landsat data," *Remote Sensing of Environment*, vol. 112, pp. 3680–3689, Sep 2008.

[44] P. A. Townsend, A. Singh, J. R. Foster, N. J. Rehberg, C. C. Kingdon, K. N. Esh-

leman, and S. W. Seagle, “A general landsat model to predict canopy defoliation in broadleaf deciduous forests,” *Remote Sensing of Environment*, vol. 119, pp. 255–265, Apr 2012.

[45] N. C. Coops, R. H. Waring, M. A. Wulder, and J. C. White, “Prediction and assessment of bark beetle-induced mortality of lodgepole pine using estimates of stand vigor derived from remotely sensed data,” *Remote Sensing of Environment*, vol. 113, pp. 1058–1066, May 2009.

[46] A. Anees and J. Aryal, “A statistical framework for near-real time detection of beetle infestation in pine forests using modis data,” *Geoscience and Remote Sensing Letters, IEEE*, vol. 11, pp. 1717–1721, Oct 2014.

[47] R. S. Skakun, M. A. Wulder, and S. E. Franklin, “Sensitivity of the thematic mapper enhanced wetness difference index to detect mountain pine beetle red-attack damage,” *Remote Sens. Environ.*, vol. 86, p. 433, Aug 2003.

[48] M. A. Wulder, J. C. White, N. C. Coops, and C. R. Butson, “Multi-temporal analysis of high spatial resolution imagery for disturbance monitoring,” *Remote Sensing of Environment*, vol. 112, pp. 2729–2740, Jun 2008.

[49] L. Eklundh, T. Johansson, and S. Solberg, “Mapping insect defoliation in scots pine with modis time-series data,” *Remote Sensing of Environment*, vol. 113, pp. 1566–1573, Jul 2009.

[50] M. Jakubauskas, D. Legates, and J. Kastens, “Harmonic analysis of time-series AVHRR NDVI data,” *Photogrammetric Engineering and Remote Sensing*, vol. 67, pp. 461–470, 2001.

[51] R. I. Negron Juarez and W. T. Liu, “Fft analysis on ndvi annual cycle and climatic regionality in northeast brazil,” *Int. J. Climatol.*, vol. 21, pp. 1803–1820, Nov 2001.

[52] A. Anees, J. Olivier, M. O’Rielly, and J. Aryal, “Detecting beetle infestations in

pine forests using modis ndvi time-series data," in *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International*, pp. 3329–3332, July 2013.

[53] C. J. Tucker, "Red and photographic infrared linear combinations for monitoring vegetation," *Remote Sens. Environ.*, vol. 8, pp. 127–150, May 1979.

[54] E. R. Hunt and B. N. Rock, "Detection of changes in leaf water content using near- and middle-infrared reflectances," *Remote Sens. Environ.*, vol. 30, pp. 43–54, Oct 1989.

[55] B. Gao, "Ndwi-normalized difference water index for remote sensing of vegetation liquid water from space," *Remote Sens. Environ.*, vol. 58, pp. 257–266, Dec 1996.

[56] A. A. Gitelson, Y. J. Kaufman, and M. N. Merzlyak, "Use of a green channel in remote sensing of global vegetation from eos-modis," *Remote Sensing of Environment*, vol. 58, pp. 289–298, Dec 1996.

[57] A. A. Gitelson, "Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation," *J. Plant Physiol.*, vol. 161, pp. 165–173, Feb 2004.

[58] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *J. Soc. Ind. Appl. Math.*, vol. 11, pp. 431–441, 1963.

[59] C. T. Kelley, *Iterative Methods for Optimization*, vol. SIAM, ch. 2, pp. 22 – 25. SIAM, 1999.

[60] R. M. Gray and L. D. Davisson, *An Introduction to Statistical Signal Processing*. New York, NY, USA: Cambridge University Press, 1st ed., 2010.

[61] W. Kleynhans, "Detecting land-cover change using modis time-series data," *Ph.D. Thesis, University of Pretoria, South Africa.*, 2011.

[62] "http://www.for.gov.bc.ca/ftp/hre/external/publish/web/bcmpb/year9/bcmpb.v9.2011kill.pdf."

[63] “http://www.for.gov.bc.ca/hfp/health/overview/mpb_h-istory.htm.”

[64] “http://foresthealth.fs.usda.gov/portal/flex/ids.”

[65] “http://landweb.nascom.nasa.gov/cgi-bin/developer/tilemap.cgi.”

3 A Statistical Framework for Near-Real Time Detection of Beetle Infestation in Pine Forests Using MODIS Data[1]

Overview

Beetle infestations have caused significant damage to the pine forest in North America. Early detection of beetle infestation in near-real time is crucial, in order to take appropriate steps to control the damage. In this letter, we consider near-real time detection of beetle infestation in North American pine forests using high temporal resolution, coarse spatial resolution MODIS (8-days 500m) satellite data. We show that the parameter sequence of a stationary vegetation index time-series, derived by fitting an underlying triply modulated cosine model over a sliding window using Nonlinear Least Squares (NLS), resembles a martingale sequence. The advantage of such properties of the parameter sequence is that standard martingale central limit (MCLT) theorem and well-known Gaussian distribution

[1]This chapter has been published as a journal article in IEEE Geoscience and Remote Sensing Letters (IEEE GRSL).

statistics can be used effectively to detect any non-stationarity in the vegetation index time-series with high accuracy. The proposed method exploits these properties of the parameter time-series , and hence does not require threshold tuning. The threshold is selected based on well-documented procedure of z-value selection from table of Gaussian distribution, depending upon the percentage of the distribution considered as outlier. The proposed framework is tested on different vegetation index datasets derived from MODIS 8-days 500 m image time-series of beetle infestations of North America. The results show that the proposed framework can detect non-stationarities in the vegetation index time-series accurately, and performs the best on Red Green Index (RGI).

3.1 Introduction

Change in vegetation occurs due to both the natural and human activities [1]. Automated land cover change detection using remotely sensed data is a topic of ongoing research [2–14]. Supervised change detection methods are attractive but the availability of ground truth data can be an issue in many practical cases [8, 14], hence unsupervised methods are popular [7]. In some change detection problems historical data is available which gives information about the normal or no-change class/behavior but, no information is available about the change class. In such cases, semi-supervised methods can be utilized. Semi-supervised methods based on statistical models may provide solutions for such problems. Of specific interest in this letter is the detection of gradual change in forests due to insect infestation, specifically the North American beetle. The North American beetle has caused significant damage to the pine forests over the past few decades [4]. Therefore, early and reliable detection of the beetle infestation is important.

Martingale theory and MCLT [15, 16] are well known in the fields of statistics. Many studies in different fields e.g. [17] have utilized martingale theory in change detection,

but to the best of our knowledge, it has not yet been investigated by the remote sensing community, specifically in land cover change detection. The advantage of using martingale theory and MCLT in change detection framework is that the threshold selection becomes an easy task and does not need tuning based on validation sets, which is desirable in any change detection framework because optimal threshold selection is always a problem. The threshold can be selected through a standard and well documented procedure of z-value selection from the table of Gaussian distribution [18].

Significant amount of prior work on detection of beetle infestation exists in the literature. Most of the existing methods utilize either high resolution ($\leq$10m spatial resolution) [9, 10] or moderate resolution ($>$10m and $<$30m spatial resolution) sensor data [3, 4]. These are not available at regular intervals of time, hence many methods use either single-date or a few images per year for analysis. This may lead to erroneous results because the same forest patch may appear very different at different time points in its phenological cycle. Therefore, the temporal resolution of the data should be high enough to differentiate real change events from the natural phenological cycles [5, 7]. MODIS provides high temporal and coarse spatial resolution data, serving the need of change detection at this analysis scale. Several studies have reported impressive results for land cover change detection using MODIS data [2, 5, 7]. Some studies have used MODIS data for detection of insect-induced tree/forest mortality [12, 13], but these studies have not utilized its full temporal resolution.

In this letter, we investigate and design a change detection framework based on martingale theory and MCLT. We utilize full temporal resolution of MODIS 8-days 500 m time-series data in this change detection framework to detect beetle infestation in pine forest. Therefore, use of an underlying triply modulated cosine function [6] is made to model the vegetation index time-series by fitting it over a sliding window using Nonlinear Least Squares (NLS). It is shown that the μ_t parameter of the triply modulated cosine function

resembles a martingale in properties. A statistical framework based on MCLT [15] is proposed for detecting non-stationarity (beetle infestation) in the parameter time-series. The threshold is selected by a well-known and well documented procedure [18] of deriving a z-value from the Gaussian distribution without any tuning. The proposed change detection framework is tested on different vegetation indices calculated from MODIS 8-day 500 m dataset of change and no-change examples taken from two study regions in North America (British Columbia and Colorado). These vegetation indices are: Normalized Difference Vegetation Index (NDVI) [19], Normalized Difference Infrared Index 6 (NDII6) [12], Normalized Difference Infrared Index 7 (NDII7) [12], Wide Dynamic Range Vegetation Index (WDRVI) [20], Infrared Simple Ratio 6 (ISR6) [11], Infrared Simple Ratio 7 (ISR7) [11] and Red Green Index (RGI) [9]. We also test the proposed change detection framework on simulated data with different magnitudes of noise, in order to analyze its robustness.

This letter is organized as follows. Section (3.2) explains the background theory and proposed methodology. Section (3.3) presents the results and discussions, and the last section concludes this letter.

3.2 Materials and Methods

3.2.1 Martingale Theory

Let X_t be a value of a sequence at time t and can be written as

$$X_t = X_{t-1} + r_t, \quad \text{for all } t > 0. \tag{3.1}$$

Let $\mathbf{X}_{t-1}$= $[X_0, X_1, \ldots, X_{t-1}]$ be a vector of all the previous values at time t. If

$$\mathrm{E}[X_t \mid \mathbf{X}_{t-1}] = X_{t-1}, \tag{3.2}$$

or

$$\mathrm{E}\left[(X_t - X_{t-1}) \mid \mathbf{X}_{\mathrm{t}-1}\right] = 0, . \tag{3.3}$$

where $\mathrm{E}[(\cdot) \mid \mathbf{X}_{t-1}]$ is expected value of $(\cdot)$ given all the observations upto time point $t-1$, then X_t is a martingale sequence [15, 16]. This means that for a martingale, the difference sequence (random jump from X_{t-1} to X_t) $r_t = X_t - X_{t-1}$ is a zero-mean Independent and Identically Distributed (*iid*) Random Variable (RV) i.e. $\mathrm{E}[r_t \mid \mathbf{X}_{t-1}] = 0$ [15, 16].

3.2.2 Martingale Central Limit Theorem (MCLT)

If X_t in (5.3) is a martingale then the MCLT [15, 16] states:

$$\lim_{\vartheta \to \infty} \frac{X_{\tau_\vartheta}}{\sqrt{\vartheta}} \sim \mathbb{N}(0,1) \tag{3.4}$$

where $\mathbb{N}(0,1)$ represents Gaussian distribution with mean 0 and standard deviation 1. For a very large value of quantity ϑ ($\vartheta \to \infty$), the time point τ_ϑ in (5.14) is given by

$$\tau_\vartheta = \min\left\{ t : \sum_{i=1}^{t} \sigma_i^2 \geq \vartheta \right\}. \tag{3.5}$$

The σ_i^2 in (5.15) is given by

$$\sigma_i^2 = \mathrm{E}\left[(X_i - X_{i-1})^2 \mid \mathbf{X}_{i-1} \right]. \tag{3.6}$$

In case r_t, for all $1 \leq t \leq N$, belong to the same distribution with variance $\sigma^2 = \sigma_1^2 = \sigma_2^2 \ldots = \sigma_N^2$, then the summation in (5.15) becomes $\sum_{i=1}^{t} \sigma_i^2 = t\sigma^2$. The inequality in

(5.15) is then satisfied at any time t by $\vartheta = t\sigma^2$, and (5.14) can be rewritten as

$$\lim_{t\to\infty} \frac{X_t}{\sqrt{t}\sigma} \sim \mathbb{N}(0,1). \tag{3.7}$$

It can be shown that (5.16) is equivalent to applying central limit theorem [18] on the difference sequence, r_t. In practical cases it is not realistic to know about the distribution or its population standard deviation (σ) and get to an infinitely large t. Therefore, sample standard deviation $\hat{\sigma}$ and a sufficiently large value of t are used in practice. The literature suggests that using $t \geq 100$ yields reliable results [18], which in case of MODIS 8-day data equals slightly above 2 years.

3.2.3 Change Detection Using Martingale Theory and MCLT

The NDVI time-series, V_t, can be modeled using a seasonal-trend model (triply modulate cosine function) [6, 7] as:

$$V_t = \mu_t + \alpha_t \cos(2\pi f t + \phi_t) + \varepsilon_t \tag{3.8}$$

where μ_t, α_t, ϕ_t and ε_t are time varying mean, amplitude, phase and noise values, respectively, and $f = \frac{1}{T}$ is the seasonal frequency with time period (cycle length) T. Let ε_t is a zero-mean *i.i.d* random variable with standard deviation $\sigma_t = \sigma$ for all $t = 1, 2, \ldots, N$. If we take a sliding window of size T and apply a linear operator Φ on it, such that:

$$d_t = \Phi([V_{t-T+1}, V_{t-T+2}, \ldots, V_t]) = \sum_{i=t-T+1}^{t} \Phi(V_i), \tag{3.9}$$

for all $T \leq t \leq N$ and

$$\Phi(0) = 0, \tag{3.10}$$

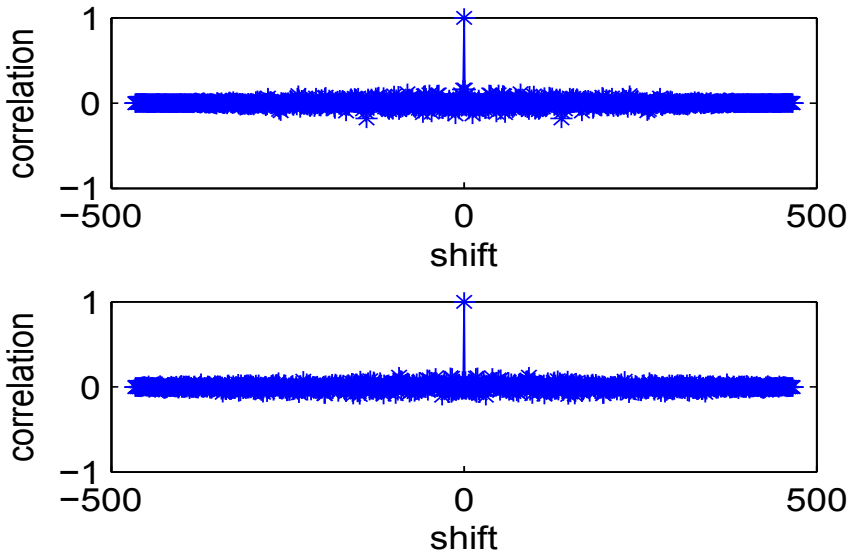

Figure 3.1: Normalized auto-correlation plot of the difference sequence of μ_t parameter sequence, derived from a real-world vegetation index time-series using NLS (top). Normalized auto-correlation plot of the difference sequence of a computer generated genuine martingale sequence (bottom). The horizontal axes represent the amount of shift in the sequence against itself.

then

$$d_t - d_{t-1} = \begin{array}{c} \left(\sum_{i=t-T+1}^{t-1} \Phi(V_i) + \Phi(V_t)\right) - \\ \left(\Phi(V_{t-T}) + \sum_{i=t-T+1}^{t-1} \Phi(V_i)\right), \end{array} \tag{3.11}$$

$$\implies d_t - d_{t-1} = \Phi(V_t) - \Phi(V_{t-T}). \tag{3.12}$$

Let $\mathbf{d}_{t-1} = [d_1, d_2, \ldots, d_{t-1}]$, then the conditional expectation of $d_t - d_{t-1}$ can be written as:

$$\mathrm{E}[(d_t - d_{t-1})|\mathbf{d}_{t-1}] = \begin{array}{c} \Phi(\mathrm{E}[V_t]|\mathbf{d}_{t-1}) - \\ \Phi(\mathrm{E}[V_{t-T}]|\mathbf{d}_{t-1}), \end{array} \tag{3.13}$$

$$\implies \mathrm{E}[(d_t - d_{t-1})|\mathbf{d}_{t-1}] = \Phi(\mathrm{E}[V_t - V_{t-T}]|\mathbf{d}_{t-1}). \tag{3.14}$$

For a perfectly stationary and noise-free time-series with periodicity T of the seasonal cycle, the values exactly T points (one-period) apart are equal i.e. $\mu_t + \alpha_t \cos(2\pi f t + \phi_t) = \mu_{t-T} + \alpha_{t-T} \cos(2\pi f (t-T) + \phi_{t-T})$. So, in case of noisy stationary time-series,

the difference, between the two points exactly one time-period apart, reduces down to the difference between the two noise values. Then (3.14) can be written as

$$\mathrm{E}[(d_t - d_{t-1})|\mathbf{d}_{t-1}] = \Phi(\mathrm{E}[\varepsilon_t|\mathbf{d}_{t-1}] - \mathrm{E}[\varepsilon_{t-T}|\mathbf{d}_{t-1}]), \tag{3.15}$$

$$\implies \mathrm{E}[(d_t - d_{t-1})|\mathbf{d}_{t-1}] = 0, \tag{3.16}$$

as ε_t is a zero-mean *iid* random variable.

$$\implies \mathrm{E}[d_t|\mathbf{d}_{t-1}] = \mathrm{E}[d_{t-1}|\mathbf{d}_{t-1}], \tag{3.17}$$

$$\implies \mathrm{E}[d_t|\mathbf{d}_{t-1}] = d_{t-1}. \tag{3.18}$$

By recalling definition of a martingale (5.8 and 5.9), (3.16) and (3.18) suggest that d_t resembles a martingale sequence.

The above mathematical treatment suggests that d_t derived from a perfectly stationary seasonal-trend time-series satisfies (5.16) and hence the MCLT. The hypothesis here is that any non-stationarity in the time-series violates (3.16) and (3.18), and d_t will no longer remain close to a martingale sequence. The non-stationarity changes the statistics of the d_t, which will make it violate (5.16). This can be detected by using Gaussian distribution z-statistic. For testing this hypothesis, first the test statistic c_t is calculated according to MCLT as

$$c_t = \mid \frac{d_t}{\sqrt{t}\hat{\sigma}} \mid \quad \forall t \geq T \tag{3.19}$$

where $\hat{\sigma}$ is the sample standard deviation of the difference sequence $[(d_i - d_{i-1}) : i = T, T+1, \ldots t]$. Then, the change alarm can be formulated as:

$$a_t = \begin{cases} 1 & \text{if } c_\mathrm{t} \geq \lambda \\ 0 & \text{otherwise} \end{cases} \tag{3.20}$$

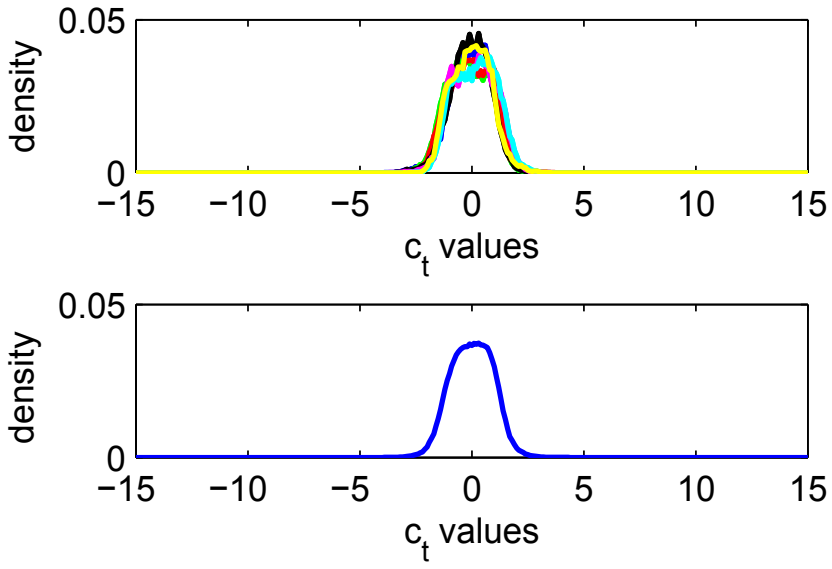

Figure 3.2: Distributions of test statistics c_t calculated according to MCLT using (3.19) from stationary time-series of 7 different vegetation indices (top). Combined distributions of all the c_t values calculated from all the vegetation index time-series, mean=0.0012, standard deviation = 0.983 (bottom).

where λ is the threshold. Initially it is assumed that the current value of d_t is coming from a stationary part, and c_t, for sufficiently large t, belongs to $\mathbb{N}(0,1)$. This assumption is tested in (3.20) by comparing c_t with the threshold λ. The threshold λ corresponds to the boundary of the Gaussian distribution, beyond which, all the values are considered as outliers and representing non-stationarity. Therefore, the alarm a_t is raised for the values of time-series, for which c_t lies beyond λ. The value of λ can be selected from the table of Gaussian distribution, through a very well known and straight forward procedure [18].

3.2.4 Beetle Infestation Detection

The above framework can be used for detecting beetle infestation in vegetation index time-series. From [3] and [21], we know that the type of change induced by the beetle infestation affects the trend component of the vegetation index time-series. There-

fore, μ_t parameter of (5.17) can be used in detecting such changes. The parameter μ_t can be derived by a T-point moving average filter, or by fitting the triply modulated cosine model over a sliding window of size T using NLS. It can be shown that μ_t satisfies (3.16) and (3.18) well enough and can be used in the above framework for change detection as d_t. The goodness of the above change detection framework for μ_t depends upon how closely μ_t sequence resembles a martingale or, in other words, how well μ_t satisfies (5.8), (5.9) and (5.14), when derived from a stationary vegetation index time-series. This can be verified by analyzing the independence of the difference sequence $[(\mu_i - \mu_{i-1}) : i = T, T+1, \ldots, t]$ (recall Section 3.2.1) i.e. its auto-correlation. Figure 3.1 shows that auto-correlation plot of the difference sequence of μ_t (top), derived from a stationary vegetation index time-series of a pine forest, is close to the auto-correlation plot of a difference sequence extracted from a genuine martingale (bottom) generated by the computer according to the theory explained in Section 3.2.1. The figure verifies that both the difference sequences are nearly *i.i.d* i.e. virtually no correlation, because there is a significant difference between the 0th (self correlation/perfect overlap) and the first correlation values (correlation value when the sequence is shifted by 1) in both the cases. Another way to verify the above arguments is to compute a large number of test statistics c_t from μ_t of the stationary seasonal-trend time-series using (3.19) and analyze their distribution. Figure 3.2 (top) shows these distributions for 7 different vegetation index time-series (1×10^5 values each). All the distributions are close to the Gaussian distribution, and have slight differences from each other, which is expected in case of real-world data. The combined distribution of these 7 distributions is shown in Figure 3.2 (bottom). Both the figures verify that the μ_t time-series, derived by fitting the triply modulated cosine model to a window of size T sliding across a stationary seasonal-trend vegetation index time-series with time period T, satisfies (5.8), (5.9) and (5.14). Therefore, the above framework for change detection can be used with confidence for detecting

Table 3.1: Results of the proposed framework on real-world beetle infestation data for all the seven indices. TP = True Positives, TN = True Negatives, Acc. = overall Accuracy and MD = Mean Detection Delay from the reference time point, $t = 230$.

Vegetation Index	TP (%)	TN (%)	Acc. (%)	MD (×8 days)
RGI	88.2	95.0	91.6	70.2
ISR6	82.8	98.0	90.4	111.2
ISR7	87.3	99.0	93.2	109.0
NDVI	78.0	92.0	85.0	90.8
WDRVI	84.2	88.0	86.1	83.22
NDII6	80.3	96.0	88.1	110.5
NDII7	87.3	97.5	92.4	108.6

beetle infestation in the μ_t time-series derived from the vegetation index time-series of the pine forests. The slight variations in the distributions plotted in Figure 3.2 can be catered for by modifying the change alarm in (3.20) as:

$$a_t = \begin{cases} 1 & \text{if } c_{\mathrm{t}} \geq \lambda\varsigma \\ 0 & \text{otherwise,} \end{cases} \tag{3.21}$$

where ς is the standard deviation of $C = [c_t : M \leq t \leq L,\ M > T]$. The variable M is the value of t which can be considered as *sufficiently large* in (5.16) and (3.19), and L is the length of the history/reference period which is known to have no changes or non-stationarities.

3.3 Results and Discussion

We collected 355 change and 355 no-change pixels of MODIS MCD43A4.005 product (8-day 500 m) from Colorado and British Columbia, North America, and prepared datasets of different vegetation indices.We tested the proposed change detection framework on these

datasets. All the examples were taken from the regions where there were no infestations recorded till the end of year 2005. Therefore, we used length of reference period equal to 230 i.e. $L = 230$ in (3.21), which corresponds to 5 years of MODIS 8-day data (2001-2005). The value of threshold we used in (3.21) was $\lambda = 3$, which was selected from the table of Gaussian distribution, corresponding to the Cumulative Distribution Function (CDF) value of 99.87%. This means that the values of c_t in (3.19) derived from stationary time-series has probability of only 0.13% to be greater than 3, and is considered as outlier. This probability increases when the non-stationarity is introduced in the time-series, and the change alarm is raised. The results of the proposed framework for all the 7 vegetation indices are summarized in Table 3.1. The general trend in the results is that the vegetation indices based on band 7 of MODIS (ISR7 and NDII7) achieve better accuracy than the rest of their counterparts based on other bands. However, the vegetation index (RGI), which is based on two such bands (Red and Green) which are sensitive to vegetation dynamics, achieves the fastest detection with the highest probability of detection. The overall accuracy is the highest in case of ISR7 but, the comparative mean detection delay from the reference time point, $t = 230$, is very high. Therefore, RGI suits near-real time detection the best because it achieves significant improvement in terms of detection delay, at the expense of only 1.6% (93.2% - 91.6%) of overall accuracy (see Table 3.1). It should be noted that the Mean Detection Delays (MD) used in this text are calculated from a common reference point of $t = 230$ i.e. end of the year 2005, and their units are 8 days.

Although the threshold λ can be selected from table of Gaussian distribution and does not need tuning, still we tested the proposed change detection framework on RGI dataset on different threshold values (350 values, from 1 to 8 with step size of 0.02) for analysis. The results are shown in Figure 3.3. The figure (top) shows that the optimal threshold lies between 2.8 and 3.4, which correspond to 99.74% and 99.97% CDF values of Gaussian distribution, respectively. By increasing the threshold further, the true positive rate

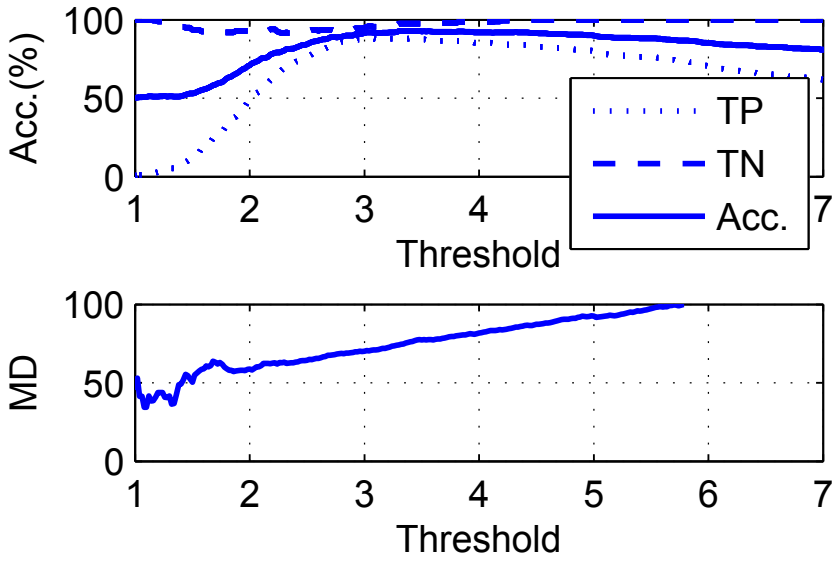

Figure 3.3: Results of the proposed framework on RGI beetle infestation data for different threshold values. Threshold vs. True Positives (TP %), True Negatives (TN %) and Accuracy (Acc. %) (top). Threshold vs. Mean Detection Delay from the reference time point, $t = 230$ (MD $\times$ 8 days) (bottom).

drops down and the detection delay increases. Note that the threshold in the proposed framework does not depend upon the type of dataset because there is an implicit standardization of the parameter time-series in (3.19), and every c_t value is the candidate of the same Gaussian distribution, regardless of the vegetation index used. This can be seen in Figure 3.2 (top) as well, where all the distributions, corresponding to stationary time-series of different vegetation indices, have near the same distributions. Therefore, the threshold used here will work for any type of dataset, as long as the time-series satisfy the conditions of the proposed framework.

In the real world dataset, we are not sure when exactly the infestation occurred, after $t = 230$, in any particular time-series. Therefore, we calculated the detection delays from the same reference point ($t = 230$) for all the vegetation indices, in order to compare their performance. To analyze the sensitivity and robustness of the proposed framework against

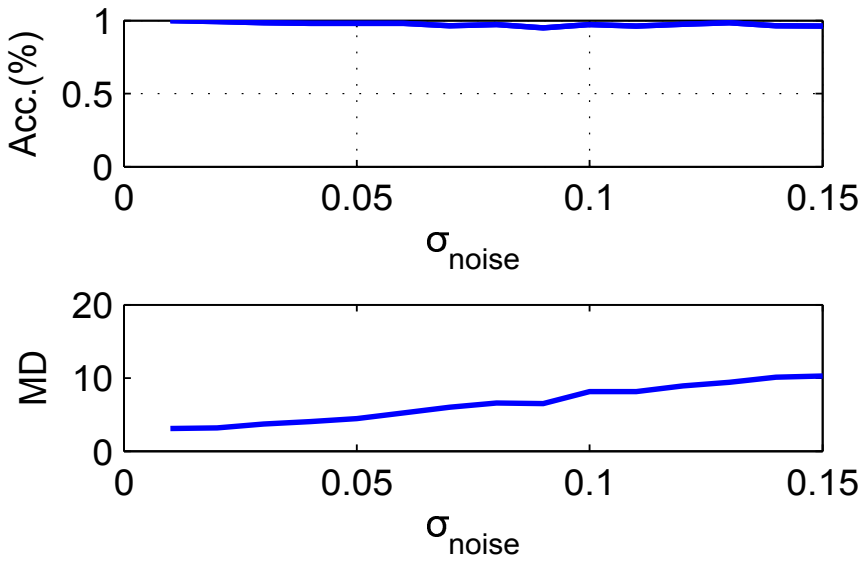

Figure 3.4: Results of the proposed framework on simulated dataset with different standard deviations of noise (σ_{noise}). Noise standard deviation (σ_{noise}) vs Accuracy (Acc.) (top). Noise standard deviation vs. Mean Detection Delay (MD) (bottom).

different magnitudes of noise, we simulated a dataset of 400 change and 400 no-change seasonal-trend time-series, each of length 850, with 15 different magnitudes of noise, following the procedure used in [2]. The change introduced in the change dataset has a magnitude of 0.3, and its location in every time-series was recorded. The results of the proposed framework for this dataset are plotted in Figure 3.4. The probability of detection does not drop much and remains around 99%, while increasing the magnitude (standard deviation) of the introduced noise. However, the detection delay increases slowly with the increase in noise. For example, the detection probability remains around 99% even if the magnitude (standard deviation) of the noise is increased from 0.1 to as high as 0.15, but the detection delay increases from 7 to 10.

3.4 Conclusions

We investigated and designed a statistical framework based on martingale theory and MCLT, for near-real time detection of beetle infestation in pine forests. The proposed method makes use of well-known martingale theory and MCLT [15, 16] for detecting non-stationarity (beetle infestation) in near-real time and hence does not need threshold tuning. The results of the proposed framework suggest the martingale theory and MCLT can be applied effectively for robust change detection, specifically detection of beetle infestation, in MODIS 8-day 500 m time-series. The best performance is achieved when RGI data is used. It is also noted that the triply modulated cosine function, originally proposed for MODIS NDVI time-series, can be used to effectively model other vegetation indices as well.

Bibliography

[1] R. Watson, I. Noble, B. Bolin, N. Ravindranath, D. Verardo, and D. Dokken, *Land Use, Land-Use Change and Forestry*. Cambridge, UK: Cambridge University Press, 2000.

[2] J. Verbesselt, A. Zeileis, and M. Herold, "Near real-time disturbance detection using satellite image time series," *Remote Sens. Environ.*, vol. 123, pp. 98–108, Aug 2012.

[3] G. W. Meigs, R. E. Kennedy, and W. B. Cohen, "A landsat time series approach to characterize bark beetle and defoliator impacts on tree mortality and surface fuels in conifer forests," *Remote Sens. Environ.*, vol. 115, pp. 3707–3718, Dec 2011.

[4] A. J. Meddens, J. A. Hicke, L. A. Vierling, and A. T. Hudak, "Evaluating methods to detect bark beetle-caused tree mortality using single-date and multi-date landsat imagery," *Remote Sens. Environ.*, vol. 132, pp. 49–58, May 2013.

[5] R. S. Lunetta, J. F. Knight, J. Ediriwickrema, J. G. Lyon, and L. D. Worthy, "Land-cover change detection using multi-temporal modis ndvi data," *Remote Sens. Environ.*, vol. 105, pp. 142–154, Nov 2006.

[6] W. Kleynhans, J. C. Olivier, K. J. Wessels, F. van den Bergh, B. P. Salmon, and K. C. Steenkamp, "Improving land cover class separation using an extended kalman filter on modis ndvi time-series data," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, pp. 381–385, Apr 2010.

[7] W. Kleynhans, J. C. Olivier, K. J. Wessels, B. P. Salmon, F. van den Bergh, and K. Steenkamp, "Detecting land cover change using an extended kalman filter on modis ndvi time-series data," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, pp. 507–511, May 2011.

[8] M. C. Hansen, R. S. Defries, J. R. G. Townshend, and R. Sohlberg, "Global land cover classification at 1 km spatial resolution using a classification tree approach," *Int. J. Remote Sens.*, vol. 21, pp. 1331–1364, Apr 2000.

[9] N. C. Coops, M. Johnson, M. A. Wulder, and J. C. White, "Assessment of quickbird high spatial resolution imagery to detect red attack damage due to mountain pine beetle infestation," *Remote Sens. Environ.*, vol. 103, pp. 67–80, Jul 2006.

[10] M. A. Wulder, J. C. White, N. C. Coops, and C. R. Butson, "Multi-temporal analysis of high spatial resolution imagery for disturbance monitoring," *Remote Sens. Environ.*, vol. 112, pp. 2729–2740, Jun 2008.

[11] R. H. Fraser and R. Latifovic, "Mapping insect-induced tree defoliation and mortality using coarse spatial resolution satellite imagery," *Int. J. Remote Sens.*, vol. 26, pp. 193–200, Jan 2005.

[12] K. DEBEURS and P. TOWNSEND, "Estimating the effect of gypsy moth defoliation using modis," *Remote Sens. Environ.*, vol. 112, pp. 3983–3990, Oct 2008.

[13] L. Eklundh, T. Johansson, and S. Solberg, "Mapping insect defoliation in scots pine with modis time-series data," *Remote Sens. Environ.*, vol. 113, pp. 1566–1573, Jul 2009.

[14] X. Zhan, R. Sohlberg, J. Townshend, C. DiMiceli, M. Carroll, J. Eastman, M. Hansen, and R. DeFries, "Detection of land cover changes using modis 250 m data," *Remote Sens. Environ.*, vol. 83, pp. 336–350, Nov 2002.

[15] B. M. Brown, "Martingale central limit theorems," *The Ann. of Math. Stats.*, vol. 42, pp. 59–66, Feb 1971.

[16] P. Hall and C. C. Heyde, *Martingale limit theory and its application*. Academic, 1980.

[17] X. Z. Kong, Y. X. Bi, and D. H. Glass, *A Geometric Moving Average Martingale method for detecting changes in data streams*, ch. 6, pp. 79–92. Springer London, 2012.

[18] D. Montgomery, *Introduction to statistical quality control*. New York, USA: Wiley, 3rd ed., 1997.

[19] C. J. Tucker, "Red and photographic infrared linear combinations for monitoring vegetation," *Remote Sens. Environ.*, vol. 8, pp. 127–150, May 1979.

[20] A. A. Gitelson, "Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation," *J. Plant Physiol.*, vol. 161, pp. 165–173, Feb 2004.

[21] A. Anees, J. C. Olivier, M. O'Rielly, and J. Aryal, "Detecting beetle infestations in pine forests using modis ndvi time-series data," *IEEE Int. Geosci. And Remote Sens. Symp. (IGARSS)*, pp. 3329– 3332, Jul. 2013.

4 A Relative Density Ratio Based Framework for Detection of Land Cover Changes in MODIS NDVI Time-Series[1]

Overview

To improve statistical approaches for near real-time land cover change detection in non-Gaussian time series data, we propose a supervised land cover change detection framework in which a MODIS NDVI time-series is modeled as a triply modulated cosine function using the Extended Kalman Filter and the trend parameter of the triply modulated cosine function is used to derive Repeated Sequential Probability Ratio Test (RSPRT) statistics. The statistics are based on relative density ratios estimated directly from the training set by a Relative Unconstrained Least Squares Importance Fitting (RULSIF) algorithm, unlike traditional likelihood ratio based test statistics. We test the framework on simulated, synthetic and real-world beetle infestation data sets, and show that using esti-

[1] This chapter has been published as a journal article in IEEE **J**ournal of **S**elected **T**opics in **A**pplied Earth Observation and **R**emote **S**ensing (IEEE JSTARS).

mated relative density ratios, instead of assuming the individual density functions to be Gaussian or approximating them with Gaussian Kernels, in the RSPRT statistics achieves better performance in terms of accuracy and detection delay. We verify the efficiency of the proposed approach by comparing its performance with three existing methods on all the three data sets under consideration in this study. We also propose a simple heuristic technique that tunes the threshold efficiently in difficult cases of near real-time change detection, when we need to take three performance indices, namely False Positives, False Negatives and Mean detection Delay, into account simultaneously.

4.1 Introduction

Land cover change detection research has seen significant recent contributions [1–19]. However, every proposed framework has its limitations on global scale due to particularity of the task at hand and the circumstances under which it is developed. Hence, no single framework is optimal in a wide range of scenarios simultaneously. Therefore, an efficient change detection framework is always in demand for a particular task and circumstances under consideration. In this study, our main focus is on the changes induced by beetle infestations in pine forests, which is one of the major causes of land cover changes in North America [2, 3]. We utilize coarse spatial resolution MODIS data due to its free-of-cost availability and high temporal resolution.

Many studies have been published which utilize coarse spatial resolution data in addressing land cover change detection [10–18, 20]. Some of these studies propose methods which are not designed for detecting changes sequentially in near real-time [11, 12, 15, 17, 18, 20]. In order to be able to mitigate the factors which are causing unwanted changes, early detection is crucial [10, 14, 16]. Therefore, considerable importance has been given to statistical approaches for near real-time land cover change detection over the recent

past [10, 13, 14, 16, 21, 22].

The term “near real-time” theoretically means that the algorithm can detect a change event with a small delay (in terms of number of observations) after the time point at which it has actually occurred, using only current and past observations. But, in remote sensing, its meaning is relative, depending upon the type of application i.e. types of changes being targeted and the data being used. Some changes are gradual and any change detection method may take a considerable number of observations before detecting them, still the methods are termed as near real-time. For example, the studies published in [10, 14, 16] introduced near real-time methods for detecting changes in MODIS time series data. The number of observations required by these methods before detecting changes may sound non-real time, but because of the type of changes addressed in these studies i.e. beetle infestations (slow and gradual) they are declared as near real-time.

Some existing methods [13, 20] derive test statistics from the raw vegetation index time series or the error between the model and observed time series [10], but it was suggested in [14, 16] that calculating change metrics from the parameter time series (time varying parameters of the fitted model) achieve better performance. There are two main issues when we consider these statistical approaches for near real-time land cover change detection. First, most of them either assume that the underlying density functions under null or alternative hypothesis are Gaussian [10, 14, 16], or they use Gaussian kernels to estimate the individual density functions [13]. The real world data may often be far from being Gaussian. This results in errors in estimated or assumed underlying distributions, which achieve suboptimal performance [23, 24]. Secondly, commonly two performance indices namely, False Positives (FP) and False Negatives (TN), are considered while evaluating the change detection methods [13]. Evaluation of a near real-time method can be based on three performance indices, namely, False Positives (FP), False Negatives (FN) and Mean detection Delay (MD), instead of only FP and FN as in the case of offline change detection

methods [13, 14]. Increasing threshold normally reduces FP rate, but increases MD and FN, and vice versa. Therefore, finding an optimal trade off manually between acceptable values of the three performance indices becomes challenging. Please note that the readers should not confuse MD with the computational time or computational complexity of the algorithm here. MD is the average detection delay calculated in terms of the number of time points (observations) between the actual point of change in the ground truth data (or selected reference point) and the point where change alarm is raised. Therefore, its units depends upon the temporal resolution of the data under consideration, which may vary from application to application. We use generic units (observations or time-points) for MD in this manuscript, which can be interpreted easily in different applications, according to the temporal resolution of the time series data being used.

It was argued in [23–27] that approximating individual densities or assuming the individual densities to be Gaussian leads to more errors. However, when the test statistics are based on density ratios we can avoid estimating individual densities or assuming them to be Gaussian as shown in [23, 24]. It was shown in [23, 25, 26] that direct density ratio estimation performs better than estimating individual densities using Gaussian Kernels or assuming them to follow Gaussanity. A number of studies have proposed different methods for direct density ratio estimation, e.g. kernel mean matching [28], the logistic regression method [29] and Kullback-Leibler importance estimation procedure (KLIEP) [25]. KLIEP was shown to be promising in change detection framework [23]. A more recent algorithm in this regard, namely Unconstrained Least Squares Importance Fitting (ULSIF), was proposed in [26] and shown to have optimal non-parametric convergence rate [30], optimal numerical stability [31] and higher robustness than KLIEP [32]. However, [27] reported a potential weakness of density ratio based approaches that density ratios can be unbounded, and proposed Relative Unconstrained Least Squares Importance Fitting (RULSIF) algorithm which uses relative density ratios that are always bounded.

RULSIF was shown to achieve better estimation and non-parametric convergence than ULSIF. Although the existing remote sensing literature on land cover change detection contains several methods based on traditional likelihood ratios with individual densities either assumed to be Gaussian, or estimated using Gaussian kernels, there is no study to our knowledge that has exploited the usefulness of relative density ratio estimation as proposed in [24, 27].

In this study, one of our main aims is to highlight and re-emphasize the usefulness of relative density ratios in remote sensing applications. We investigate the advantages of using relative density ratio estimation in supervised near real-time classification of change and no-change events within the MODIS NDVI time series. Our proposed change detection framework models MODIS 8-days 500m NDVI time-series by a triply modulated cosine function [17, 18] and uses the Extended Kalman Filter (EKF) [33] to derive its time varying parameters. As suggested in [1, 14, 15], changes which effect the trend of the signal can be captured in the trend parameter of the triply modulated cosine function. Therefore, our proposed framework learns relative density ratios from the trend parameters of the change and no-change training sets using the RULSIF algorithm [24]. Once the training is done, these estimated relative density ratios are used to derive RSPRT (Repeated Sequential Probability Ratio Test) [33] statistics online, which can be compared to a tuned threshold in order to detect changes in near real-time.

We also address the issue of finding an acceptable trade off between FP, FN and MD while tuning the threshold. Traditionally, performance of the any change detection method is analyzed based only on accuracy. In such cases, the threshold tuning can be done with the help of Receiver Operating Characteristics (ROC) curve or calibration curve [34]. The ROC curve helps in finding a trade off between TP and FP, and the calibration curve plots accuracy against different values of threshold. But, in our case, we have to consider MD as well, along with accuracy, which means that the optimal trade off has to be found be-

tween three performance indices, namely FN (instead of TP), FP and MD . We formulate a cost function dependent upon FP, FN, MD (or alternatively kappa-coefficient [35] and MD) and the threshold, which is minimized iteratively to find a threshold value that gives acceptable trade off between these performance indices. Finally, we compare our proposed framework with 3 recently published land cover change detection methods for land cover change detection in MODIS NDVI time series data which use either kernel density estimation to estimate the individual densities or assume them to be Gaussian, while deriving the test statistics [13]. We show that our proposed framework achieves better accuracy with lower detection delays.

The main contributions of this study are: 1) A supervised near real-time change detection framework that can detect land cover changes in MODIS NDVI time-series data quicker and with more accuracy than recently published methods. 2) Highlighting and re-emphasizing the usefulness of density/relative density ratio estimation [23, 24, 27] in the remote sensing community while showing its suitability for supervised change detection in MODIS time-series data, 3) An effective strategy for tuning the threshold automatically in near real-time scenarios when more than two performance indices have to be considered, and also in those scenarios where manual threshold selection is cumbersome e.g. cross-validation experiments.

The research questions addressed in this study are: I) Is the relative density ratio estimation a viable option for supervised change detection in MODIS time series data? II) Do RSPRT/CUSUM (CUmulative SUM) statistics [33, 36, 37], when derived from the parameter time-series, improve performance compared to when derived from the raw time-series [13]? III) Does using the relative density ratios, estimated by RULSIF [23, 24, 27] in RSPRT statistics improve the performance compared to estimating the individual densities [13] or assuming them to be Gaussian [10, 14]?

The paper is organized as follows. Section 4.2 explains the RULSIF algorithm for relative

density ratio estimation, the proposed supervised land cover change detection framework, and the proposed threshold tuning technique. Section (4.3) gives brief descriptions of the three existing methods used in this study for comparison and performance evaluation of our proposed framework. Section (4.4) explains the data sets used in this study. Section 4.5 presents the numerical results, their comparison and discussion. The last section concludes this paper.

4.2 Materials and Methods

4.2.1 Repeated Sequential Probability Ratio Test (RSPRT) with Relative Density Ratio Estimation (M1)

MODIS time series data contains seasonal variations which need to be taken into account while designing any change detection method [17, 38]. Different types of functions have been used to model MODIS vegetation index time series, in land cover change detection framework, over the recent past [10–12, 14, 16–18]. Some recent studies have argued the usefulness of triply modulated cosine function and its time varying parameters, in land cover change detection framework [14, 16–18]. In order to get trend and seasonal variations separately, we model the vegetation index time-series of a given MODIS pixel by a triply modulated cosine function as in [14, 16–18]

$$y_t = \mu_t + \alpha_t \sin(2\pi f t + \phi_t) + v_t \tag{4.1}$$

where y_t and v_t are the observation and noise value from an unknown distribution, at time $t = 1, 2, \ldots$. The above model is based on many unknown parameters, namely the frequency f, and the time varying parameters mean μ_t, amplitude α_t and phase ϕ_t. The parameter f is determined by the data being used for analysis. In our case, the MODIS

8-day 500m time-series has cycle-length of 1 year with 46 observations per year, hence $f = 1/46$. The values of μ_t, α_t and ϕ_t can be estimated from the observations y_t according to (4.1) using a non-linear estimator. As proposed in [18], EKF can be used to derive the time-varying parameters of (4.1). In EKF formulation, the model given in equation (4.1) can be written as a pair of state and measurement equations as

$$\mathbf{x}_t = \mathbf{v}(\mathbf{x}_{t-1}) + \mathbf{w}_t \tag{4.2}$$

and

$$\mathbf{y}_t = \mathrm{h}(\mathbf{x}_t) + v_t \tag{4.3}$$

where $\mathbf{x}_t = [\mu_t, \alpha_t, \phi_t]^T$ is the state vector, $\mathbf{v}$ is the relationship between the previous state and the current state, $\mathbf{w}_t$ is 3×1 vector of process noise at time t, v_t is the measurement noise at time t and, $\mathbf{h}$ is the relationship between the current state $\mathbf{x}_t$ and the predicted measurement $\mathbf{y}_t$. The EKF predicts the state vector at time t recursively [18], using the observations till time t.

Estimation of the state vector at every time point t using EKF results in time series of the parameters. The next step is to compute the change metrics/test statistics in order to classify change or no-change events. As shown in [1, 14, 15], the trend changes e.g. changes due to beetle infestations create significant impact on μ_t, hence we calculate our test statistics from μ_t. Many types of control charts exist in literature, e.g. Shewhart control charts [33, 36, 39], Moving Average control charts [33, 36], RSPRT/CUSUM control charts [33, 36, 37], Generalized Likelihood Ratio (GLR) control charts [33, 36] etc., which can be applied in deriving the test statistics. However, RSPRT/CUSUM detect small changes quicker (takes lesser number of observations or data points after the change has occurred) than rest of the control charts [33, 36]. Since the type of change we are targeting here is of gradual nature, we use RSPRT to derive test statistic S_t from μ_t time

series as

$$S_t = \begin{cases} S_{t-1} + \ln\frac{p_{H_1}(\mu_t)}{p_{H_0}(\mu_t)} & \text{if } S_{t-1} + \ln\frac{p_{H_1}(\mu_t)}{p_{H_0}(\mu_t)} > 0 \\ 0 & \text{if } S_{t-1} + \ln\frac{p_{H_1}(\mu_t)}{p_{H_0}(\mu_t)} \leq 0 \end{cases} \tag{4.4}$$

where $p_{\mathscr{H}_*}(\mu_t)$, is the likelihood of vector random variable μ_t, at time t, under hypothesis $\mathscr{H}_*$, and $S_0 = 0$. The vector random variable $\mu_t = [\mu_t, \mu_{t-1}, ..., \mu_{t-k+1}]$ in (4.4) is derived with a sliding window of length k in order to capture the relationship of μ_t with its immediate past. The value of k can be chosen by the user (normally $k \geq 10$, we used $k = 10$). The no-change and the alternate hypotheses, $\mathscr{H}_0$ and $\mathscr{H}_1$ respectively, can be defined as

$$\begin{aligned} \mathscr{H}_0 : &\quad S_t \leq \lambda \\ \mathscr{H}_1 : &\quad S_t > \lambda \end{aligned} \tag{4.5}$$

where λ is a carefully selected threshold. Equation (4.4) can be compacted as

$$S_t = (S_{t-1} + s_t)^+ \tag{4.6}$$

where $(\rho)^+ = \sup(0, \rho)$ for some value of ρ, $s_t = \ln\frac{p_{\mathscr{H}_1}(\mu_t)}{p_{\mathscr{H}_0}(\mu_t)}$. The change alarm a_t at time t can be raised according to

$$a_t = \begin{cases} 1 & \text{if } S_t > \lambda \\ 0 & \text{if } S_t \leq \lambda \end{cases} \qquad t \geq k \tag{4.7}$$

for a carefully selected threshold λ.

The likelihood ratio in RSPRT is often found either by assuming the individual density functions to be Gaussian [33] or by estimating the individual density functions using kernel Density Estimation (KDE) [13]. Both the methods can lead to suboptimal results because the real world data rarely satisfies Gaussanity condition, and density estimation too is a difficult problem to solve [23, 24, 27, 40]. Estimating the density ratios directly,

without estimating the individual distributions, is comparatively easier and achieves better performance [23, 24, 27, 40]. Although many algorithms have been used for direct density ratio estimation, e.g. KLIEP [23, 25], ULSIF [24, 26], [24, 27] suggested that RULSIF algorithm, which considers relative density ratios, achieves better estimation and non-parametric convergence. Therefore, we use relative density ratios, estimated directly from change and no-change training sets using RULSIF algorithm [27], in (4.4). Let $\mathbf{Y}_{tr}$ be the training set containing change and no-change ground truth examples. After deriving the parameters of equation (4.1) using EKF, the change and no-change training sets of the μ_t parameter denoted, respectively, by $\mathrm{Y_c} = \{\mu_i\}_{i=1}^{n}$ and $\mathrm{Y_{nc}} = \{\mu'_i\}_{i=1}^{m}$, can be formed by sliding a window of length k over change and no-change μ_t time series as $\mu_t = [\mu_t, \mu_{t-1}, ..., \mu_{t-k+1}]$ starting at $t = k$, and putting the window values, at each time point, in the respective sets. According to the RULSIF formulation the relative density ratio can be given by [24, 27]

$$r_\beta(\mu) = \frac{p(\mu)}{\beta p(\mu) + (1-\beta)p'(\mu)} = \frac{p(\mu)}{p'_\beta(\mu)} \tag{4.8}$$

where $p(\mu)$ represents the density of change samples, $p'(\mu)$ represents the density of no-change samples, μ is an arbitrary data sample and $0 \leq \beta < 1$. The β-relative density ratio $r_\beta(\mu)$ can be estimated by a kernel model as [24, 27]

$$r_\beta(\mu) \approx g(\mu;\theta) = \sum_{i=1}^{n} \theta_i K(\mu, \mu_i) \tag{4.9}$$

$$K(\mu, \mu_i) = \exp\left(-\frac{\| \mu - \mu_i \|^2}{2\sigma^2}\right) \tag{4.10}$$

where $\theta = [\theta_1, \theta_2, ..., \theta_n]$ is the parameter vector, σ(>0) is the kernel width and n is the number of change examples in the training set. The appropriate value of σ was selected as explained in [24].

Note that the complexity of problem increases with the increase in size of the training data set because the number of kernels being used and hence the number of θ parameters which need to be estimated, is equal to the number of samples in training set. This works fine for small training sets but it introduces memory and computation time issues in case of the large training sets. Therefore, we use a small but sufficient number of centers selected randomly from the training set, instead of using all the samples of the training set as centers. We adapt (4.9) as

$$r_\beta(\mu) \approx g(\mu;\theta) = \sum_{i=1}^{d} \theta_i K(\mu, \eta_i) \tag{4.11}$$

where $\{\eta_i\}_{i=1}^{d}$ is a set of d number of centers chosen randomly from the training set. As formulated in RULSIF, the squared loss between the true and estimated relative density ratios, $J(\mu)$ is given by [24, 27]

$$\begin{aligned} J(\mu) &= \tfrac{1}{2}\int p'_\beta(\mu)\left(r_\beta(\mu) - g(\mu;\theta)\right)^2 \mathrm{d}\mu \\ &= \tfrac{1}{2}\int p'_\beta(\mu)\left(r_\beta(\mu)\right)^2 \mathrm{d}\mu \\ &\quad - \int p(\mu) g(\mu;\theta)\mathrm{d}\mu \\ &\quad + \tfrac{\beta}{2}\int p(\mu)\left(g(\mu;\theta)\right)^2 \mathrm{d}\mu \\ &\quad + \tfrac{1-\beta}{2}\int p'(\mu)\left(g(\mu;\theta)\right)^2 \mathrm{d}\mu. \end{aligned} \tag{4.12}$$

The parameter vector θ can be estimated by minimizing $J(\mu)$. Ignoring the terms independent of $g(\mu;\theta)$ in (4.12), the following optimization problem is formulated according to RULSIF [27]

$$\min_{\theta\in\mathbb{R}^n}\left[\frac{1}{2}\theta^T\widehat{H}\theta - \widehat{h}^T\theta + \frac{\gamma}{2}\theta^T\theta\right] \tag{4.13}$$

$$\implies \theta = \left(\widehat{H} + \gamma I_n\right)^{-1}\widehat{h} \tag{4.14}$$

where $\widehat{H}$ is a $d \times d$ matrix, I_d is an d dimensional identity matrix, $\widehat{h}$ is a vector of length d

and $\gamma \geq 0$ is a regularization parameter. The (l,l')th element of $\widehat{H}$, for all $1 \leq l,l' \leq d$, is given by [27]

$$\widehat{H}_{l,l'} = \frac{\beta}{n}\sum_{i=1}^{n} K(\mu_i,\eta_l)K(\mu_i,\eta_{l'}) + \frac{1-\beta}{m}\sum_{j=1}^{m} K(\mu'_j,\eta_l)K(\mu'_j,\eta_{l'}) \tag{4.15}$$

and lth element of $\widehat{h}$ can be given by [27]

$$\widehat{h}_l = \frac{1}{n}\sum_{i=1}^{n} K(\mu_i,\eta_l). \tag{4.16}$$

Once the parameter vector θ has been estimated, it is used in (4.11) to estimate the relative density ratio of any μ_t which is then used in (4.4) to calculate the test statistics. First we find the relative density ratio sequences for all the examples in the training set. The threshold λ is then tuned using this training set. Then in the similar way relative density ratio is found at any time t in test time-series as well, and change alarm can be raised according to (4.7). Both the training and testing phases of the proposed framework have been summarized in Algorithm 4.1 and Algorithm 4.2, respectively. It is worth noting here that the authors of [13, 41] mentioned that both independent and identically distributed (i.i.d) assumptions were not met in their formulation of CUSUM. After removing the seasonality, the unchanged μ_t parameter time series has slightly reduced correlation but not enough to be considered as negligible. However, our formulation considers all the no-change samples as coming from a single distribution, unlike CUSUM formulation in [13, 41]. Therefore, the no-change samples can be considered as identically distributed, and the change is detected when this assumption is violated i.e. when a sample from a significantly different distribution is encountered.

Algorithm 4.1 training($\mathbf{Y}_{tr}$,k, β, σ , γ)

Given the training set $\mathbf{Y}_{tr}$, k, β, σ and γ

1. Derive the parameters $\mathbf{x}_t = [\mu_t, \alpha_t, \phi_t]$ for all training time-series using EKF.
2. Make separate sets for change and no-change samples, $\mathrm{Y}_c = \{\mu_i\}_{i=1}^{n}$ and $\mathrm{Y}_{nc} = \{\mu_i'\}_{i=1}^{m}$, respectively.
3. Chose d samples from Y_c randomly as kernel centers ($\eta = \{\eta_i\}_{i=1}^{d}$).
4. Use η,$\mathrm{Y}_c$and Y_{nc} in equation (4.13) to (4.16) to estimate the parameter vector θ of the equation (4.11).
5. At every time-point t of the training time-series estimate $r_\beta(\mu_t)$ using (4.11).
6. Use the estimated $r_\beta(\mu_t)$ in place of $\frac{p_{\mathscr{H}_1}(\mu_t)}{p_{\mathscr{H}_0}(\mu_t)}$ in equation (4.4) to calculate RSPRT statistic S_t.
7. Using RSPRT statistics of the whole training set, tune an optimal threshold (λ) that minimizes False Negatives (FN), False Positives (FP) and Mean Detection Delay (MD).

OUTPUT (θ , λ, η)

Algorithm 4.2 test(y_t, k, η,Y_c and σ)

Given all the observations till current time t ($\mathbf{y}_t$), k, η, Y_c and σ

1. Derive the parameters $\mathbf{x}_t = [\mu_t, \alpha_t, \phi_t]$ of (4.1) using EKF.
2. Derive a test sample $\mu_t = [\mu_t, \mu_{t-1}, ..., \mu_{t-k+1}]^T$.
3. Estimate $r_\beta(\mu_t)$ using (4.11).
4. Use the estimated $r_\beta(\mu_t)$ in place of $\frac{p_{\mathscr{H}_1}(\mu_t)}{p_{\mathscr{H}_0}(\mu_t)}$ in equation (4.4) to calculate RSPRT statistic S_t.
5. Use (4.7) to evaluate the change alarm a_t.

OUTPUT (a_t)

4.2.2 Automatic Threshold Tuning

Tuning a threshold manually on training data set, while taking care of 3 performance indices i.e. FN, FP and MD, is a challenging task. Here we formulate a simple, yet effective constrained optimization problem, which will yields a tuned threshold. Since all the three performance indicators are desired to be as low as possible, an over all cost function based including the effect of all of them can be designed, which can then be minimized. A simple choice can be Euclidean norm written as

$$\mathscr{L}_q = \sqrt{(FP_q)^2 + (FN_q)^2 + (\psi \times MD_q)^2} \tag{4.17}$$

where $\mathscr{L}_q$ is the cost at the qth iteration of the optimization algorithm and ψ is the weight that increases or decreases the dependence of the cost function/optimization on MD. A properly selected ψ also caters for the scale difference between MD and rest of the two indicators. Its value can be selected against a desired accuracy in the first run of the cross validation and kept the same throughout rest of the experiment. Please note that FP, FN and MD are derived from (4.7) which means that they are dependent on λ and hence the cost function as well. So, ideally, minimization of the cost function in (4.17) subject to $\lambda > 0$, should yield optimal value of λ for a specific value of ψ. However, at some instances, the cost function may remain (flat) unchanged with the change in the value of λ, as shown in Figure 4.1 (top). This is undesirable because the optimization algorithm may get stuck in such "flat" regions and stop prematurely. Mathematically, the flat regions in Figure 4.1 (top) violate

$$P\left\{\left|\frac{d\mathscr{L}_q}{dq}\right| > 0 \mid \mathbf{E}\left[\frac{d\mathscr{L}_q}{dq}\right] < 0\right\} = 1 \tag{4.18}$$

for all $q < N$, where N is the iteration number at which the optimization algorithm converges and $\mathscr{L}_N < \mathscr{L}_{q \neq N}$. The operators $P\{\bullet\}$ and $\mathbf{E}\{\star\}$ represent probability and expec-

tation, respectively. In order to tackle this issue we modify (4.17) slightly and introduce stochasticity in it as

$$\mathscr{L}'_q = \sqrt{(FP_q)^2 + (FN_q)^2 + (\psi \times MD_q)^2} + \varepsilon_q \tag{4.19}$$

where $\mathscr{L}'_q$ is the value of stochastic cost and ε_q is a small random number drawn from uniform or Gaussian distribution, in qth iteration. The effect of randomness/stochasticity has been shown in Figure 4.1 (bottom). It can be seen in Figure 4.1 (bottom) that (4.19) satisfies (4.18). Note that the condition in (4.18) is not strict. So, there is still a possibility that the optimization algorithm may get trapped in local minimum and stop prematurely, without converging to an optimal value of λ. Therefore, the minimization must be carried out several times (e.g. 10-20 times), each time with different initial value of λ, and the one with the lowest value of the cost function after convergence, should be selected. Our experiments with different optimization algorithms suggest that Genetic Algorithm (GA) is able to optimize (4.17) successfully.

The weight ψ varies the importance of MD in the cost function i.e. increasing ψ will cause minimization to occur at lower values of MD and vice versa. This relationship can be seen in Figure 4.2 which shows ψ vs. MD plot. Mathematically it can be written as

$$\mathbf{E}\left[\frac{dMD}{d\psi}\right] < 0. \tag{4.20}$$

An alternative for the above threshold selection strategy can be based on kappa-statistic (κ) [35, 42] and MD. Use of κ- statistic in remote sensing to measure homogeneity is somewhat controversial and there has been some criticism by some authors on the accuracy assessment based on κ-statistic, because its value depends strongly on the marginal

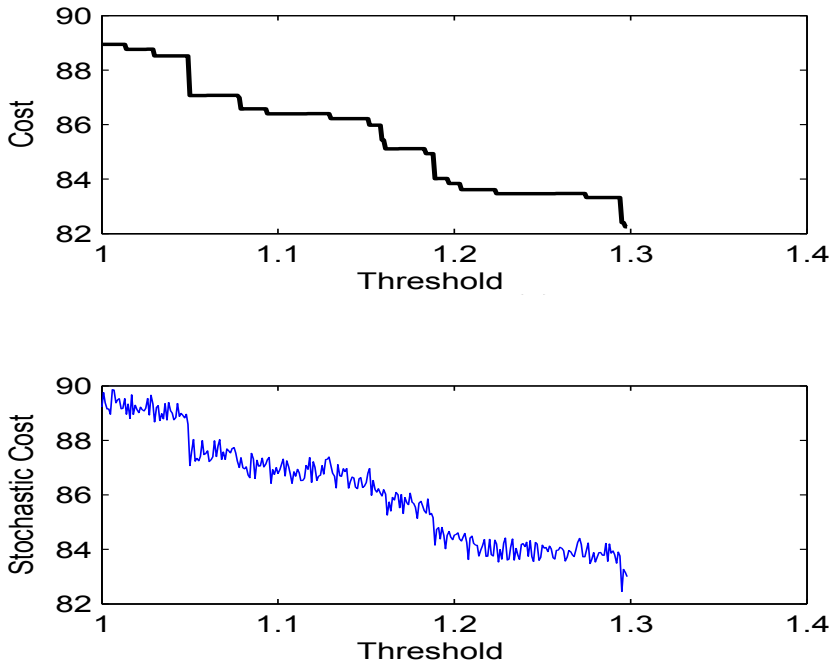

Figure 4.1: Threshold (λ) vs. Cost (top). Threshold (λ) vs. Stochastic Cost (bottom)

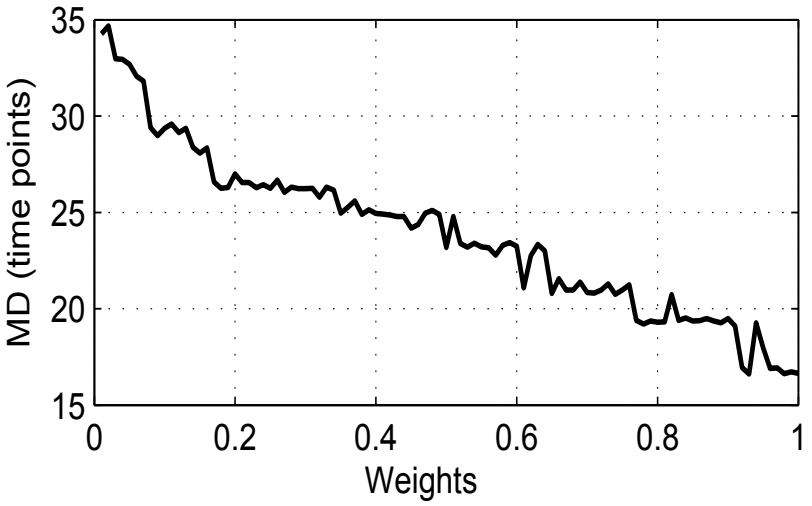

Figure 4.2: Weight (ψ) vs. *MD* plot. Increasing weights (ψ) decreases the acceptable *MD* values. The unit of MD is time points = number of time points or number of observations.

distributions [43–45]. However, it is still the most widely used statistic, hence the usage here. It can be calculated as

$$\kappa = \frac{N \times -\{(TP' + FP') \times (TP' + FN') + (TN' + FP') \times (TN' + FN')\}}{N^2 - \{(TP' + FP') \times (TP' + FN') + (TN' + FP') \times (TN' + FN')\}}, \tag{4.21}$$

where TP', TN', FP' and FN' are the true values (not percentages) of true positives, true negatives, false positives and false negatives, respectively, and $N = TN' + FN' + TP' + FP'$. The optimal threshold, corresponding to acceptable values of MD or κ-coefficient, can be selected from "κ-coefficient vs. MD" plots. This can also be achieved by minimizing a cost function based on κ and MD using genetic algorithm, subject to desired constraints on MD and/or κ, which can be easily programmed and integrated in cross validation experiments. We used this strategy in our cross validation experiments (see Section 4.5.4). The advantage of $\kappa-$statistic based threshold selection strategy is that it incorporates all the performance indices into a single coefficient and the threshold selection becomes more convenient as it can be selected from a plot of two variables. However, if there are some special constraints on TP and TN (or FP and FN) which need to be followed in a certain application, then the former threshold selection strategy is more convenient.

4.3 Existing Methods

4.3.1 Original CUSUM with Kernel Density Estimation (M2)

A supervised method, implementing CUSUM, was proposed in [13] for land cover change detection in MODIS NDVI time-series data. In this method CUSUM statistics are derived

from raw MODIS NDVI time-series. The individual density functions required for calculating likelihood ratios are derived for every time point in the year/cycle using kernel density estimation. The MODIS product used in this study has the time resolution of 46 images/ year. Therefore, using this method, estimation of total of 92 density functions (46 density functions separately for both change and no-change hypotheses) is required. Once the density functions are estimated, the likelihood ratio (density ratio) of an observation at anytime point in the cycle is calculated using the trained density functions of change and no-change for that particular time of the cycle/year. The likelihood ratio is then used to calculate the CUSUM statistic sequentially, which is then compared with a tuned threshold in order to detect any change event. Since this study uses MODIS NDVI data and the method is based on CUSUM statistics with estimated individual densities, comparison with this method will give us a good insight into advantages of model fitting and using estimated/trained relative density ratios rather than estimated/trained individual densities.

4.3.2 Near Real-Time Disturbance Detection (M3)

Another method that was published recently in [10], is the "Near Real-Time disturbance Detection in MODIS Data". In this method, MODIS NDVI time-series is modeled using a function with constant, ramp, sine and cosine terms. First, the function is fitted to the reference (no-change) period using non-linear least squares fitting and its unknown parameters are derived. Then using these parameters, the future observations are predicted using the model. The difference between the predicted and observed values gives noise time-series. The MOSUM (MOving SUM) statistics [46, 47] are then derived from the noise time-series and compared with a threshold, tuned according to functional central limit theorem [48], in order to detect any change events. Since this method uses noise time-series to derive test statistics and also Gaussanity assumption which is implicit in

central limit theorem, comparison with it will give us a good insight into benefits of using parameter time-series and also relative density ratio estimation over assuming individual densities to be Gaussian.

4.3.3 Near Real-Time Detection of Beetle Infestation (M4)

One of the most recent studies which address near real-time detection of land cover changes, specifically beetle infestations in pine forests, using MODIS NDVI data was published in [14]. In this method, non-linear least squares approach is used to fit a model to the NDVI time series and derive its time varying parameters. Based on the fact that beetle infestation effect the trend of the signal significantly, the trend component of the model is used to derive the test statistics. It is assumed that the underlying densities of the change and no-change parts of the trend component are nearly Gaussian with difference in their means. Based on this assumption, the log likelihood ratio of the value of the trend component, at any particular time point, is calculated. This log likelihood ratio is then compared to a tuned threshold in order to declare a change or no-change event. The threshold is tuned by finding a good trade off between the likelihood ratios of change and no change training sets.

4.4 Data sets

4.4.1 Simulated Data

One main problem that is often encountered in case of near real-time change detection is that the ground truth data with accurate labels/time-point of change events is hard to find [11]. Most often, partial information is known e.g. changed and unchanged pixels are known but the exact time-points at which the changes occurred in the respective

time-series are unknown, hence making the performance evaluation difficult and compromised. Many studies have used and highlighted the importance of simulated data, in which, changes are introduced at desired time-points [10–12, 14, 16, 49]. Such data can be helpful in evaluating performance, sensitivity and robustness of the method to different magnitudes of noise. We also generated a simulated data set following a similar procedure as used in [10–12, 14, 16].

First, the deterministic part or seasonal cycles were generated using asymmetric Gaussian function as proposed in [12, 14]

$$g(l) \equiv g(l;a,b,\rho_1,\rho_2) = a \times \begin{cases} \exp\left[-\frac{(l-b)^2}{\rho_1}\right], & \text{if } l > b \\ \exp\left[-\frac{(b-l)^2}{\rho_2}\right], & \text{if } l < b, \end{cases} \tag{4.22}$$

where ρ_1 and ρ_2 control the width of the left and right hand sides, whereas a and b are the amplitude and the position of the maximum or minimum with respect to the time l, respectively. We used $a = 0.7$, $b = 23 + \lfloor l/46 \rfloor \times 46$ and $\rho_1 = \rho_2 = 100$ in (4.22). The simulated time-series were generated by

$$S(l) = g(l) + \Phi(l) + \vartheta_l \tag{4.23}$$

where ϑ_l is a noise value at l, drawn from Gaussian distribution with zero mean. The $\Phi(l)$ in (4.23) is the simulated gradual change introduced in the time-series, and can be given by

$$\Phi(l) = \begin{cases} \left[(1)^{\lfloor l/\rho \rfloor} - (0)^{\lfloor l/\rho \rfloor}\right] \times \varsigma \times (l-\rho) & \text{if } l \leq \xi \\ \Phi(\xi) & \text{if } l > \xi, \end{cases} \tag{4.24}$$

where ς, ρ, and ξ are the slope, start point and end point of the introduced change, respectively. Note that simulating remotely sensed data with vegetation phenology, inter-annual variability, disturbance events and signal contamination is challenging [11]. Therefore,

testing the method on a variety of data sets is necessary.

4.4.2 Synthetic Data

Simulated data set is far from real-world data and the factors involved in it. To get as much close as possible to the real-world data, yet knowing the exact time-points of the change events, another type of data, namely synthetic data set, has also been used [13, 17, 18, 50, 51]. This data can be created from sure change and no-change parts of the real-world time-series. First, all the time-series are standardized according to the range of the no-change part and then different no-change parts are concatenated with different change parts randomly, creating a large number of time-series. So, all the time-series have natural factors involved in them as well as the exact time-points of change events are also known.

4.4.3 Real-World Beetle Infestation Data

We used yearly survey maps and shape files maintained by the United States and British Columbia forest services [52–54] to identify the areas with beetle infestations in the pine forests of Colorado, Utah (United States), and British Columbia (Canada). The regions with no beetle infestation history till the end of 2005 were selected. These regions were then marked on the Google Earth and their geographical coordinates were recorded. An online tool (MODLAND Tile Calculator) [55] was then used to identify the corresponding MODIS tile using the geographical coordinates. Once the MODIS tile was known, one MODIS image of 500 meters spatial resolution was fed into a software, namely, ENVI (version 5). Using the geographical coordinates of the marked regions, their pixel co-ordinates, in any MODIS image of 500 m spatial resolution, were found with the help of ENVI. After recording all this information, the MODIS product MCD43A4.005 was

downloaded for the desired tiles starting from January 2001 to December 2011, and the time series of the selected pixels were extracted. All the change and no-change examples from both the regions were combined and two (change and no-change) data sets were prepared.

The MODIS product MCD43A4.005 is available since 2000. It provides 500 m 8-days composite reflectance data which is Bidirectional Reflectance Distribution Function (BRDF)-adjusted for Nadir reflectance, atmospherically corrected and cloud free . The data acquired for the year 2000 had a lot of missing values, hence was discarded and the time series data was acquired from Jan. 2001 onwards. Although we did not encounter any missing values in our analysis, however rare missing values can be replaced with interpolated values.

4.5 Results, Comparison and Discussion

4.5.1 Results for Simulated Data

We generated 500 change and 500 no-change examples according to the methodology used in [10–12, 14, 16] summarized in Section 4.4.1. The seasonal cycles were generated using asymmetric Gaussian function, the change was introduced by adding a ramp of slope 0.0025 ($\varsigma = 0.0025$ in Equation 4.24) to the signals at known positions in order to replicate a gradual change, and the noise introduced into the signal was drawn randomly from the noise distribution with standard deviation of 0.08. Randomly selected 50% samples of the data set was taken as the training set and the rest 50% was taken as the test set. The purpose of this data set was to analyze the performance of the proposed framework (M1) on a data set with known change points and compare it with the performances of the existing methods. Moreover, it can also be used to analyze the robustness of the proposed

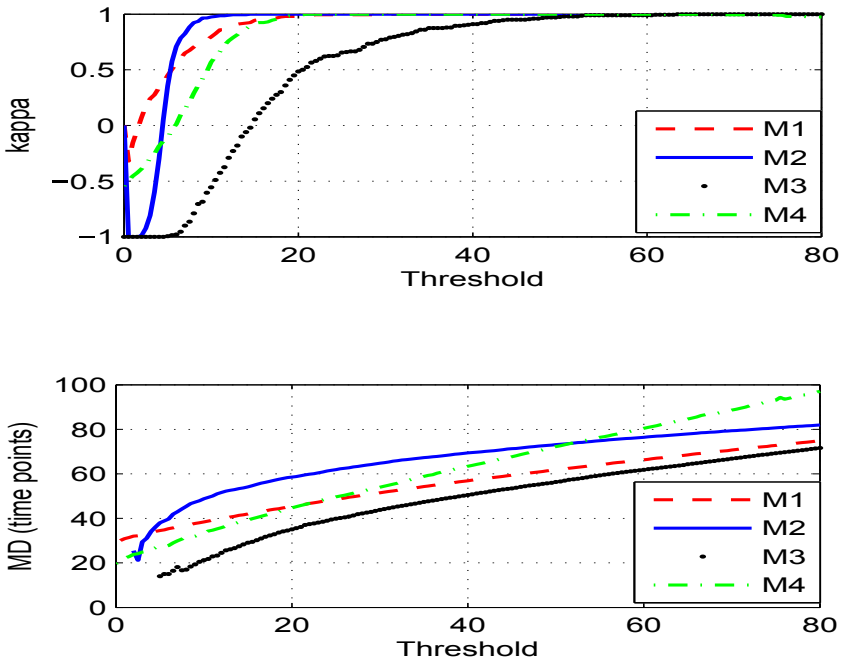

Figure 4.3: Comparison between performances of proposed (M1), original CUSUM (M2), near real-time disturbance detection (M3) and near real-time beetle infestation detection (M4) methods, on simulated data. Top: Threshold (λ) vs. κ-coefficient (top). Threshold (λ) vs. MD (bottom). The unit of MD is time points = number of time points or number of observations.

method M1 against different magnitudes (standard deviations) of noise.

All the four methods, M1-M4, were implemented keeping the training and test sets exactly the same in order to ensure fair comparison. A wide range of threshold values were used for each method in order to exploit its performance range and capabilities. The results of all the methods, M1-M4, have been summarized in Figure 4.3, Figure 4.4 and Table 4.1. Please note that all the 4 methods have different ranges of threshold values, but here we have scaled all of them to a single range of 0-80, for the sake of simplicity in comparison. The absolute values of the thresholds are not important here because we only want to graphically present the best possible performances by each of the methods considered here. Figure 4.3 (top) presents the "Threshold vs. κ-coefficient" plots of

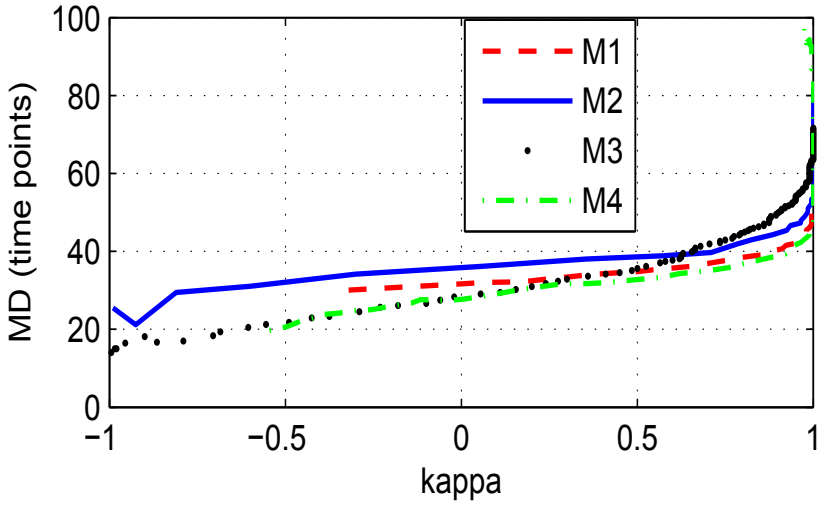

Figure 4.4: Comparison of "kappa-coefficient vs. MD" performances of proposed (M1), original CUSUM (M2), near real-time disturbance detection (M3) and near real-time beetle infestation detection (M4) methods on simulated data. The unit of MD is time points = number of time points or number of observations.

Table 4.1: Comparison between near real-time performances of our proposed method (M1) and the existing three methods, on simulated data, at acceptable true positives, true negatives and accuracy. TP = True Positive, TN = True Negative, Acc. = Overall Accuracy, MD = Mean Detection Delay, and λ = threshold value. M1= Proposed Framework, M2 = Original CUSUM Method , M3 = Near Real-Time Disturbance Detection Method(M3) and M4 = Near Real-Time Beetle Infestation Detection Method. The units of TP, TN, Acc. are "%" and that of MD is tp = number of time points or observations.

Method	TP	TN	Acc.	MD (tp)	Threshold (λ)
M1	100	98	99	45	18.50
M2	98.50	99.50	99	49	10.50
M3	100	98	99	56	05.05
M4	100	98	99	44	03.80

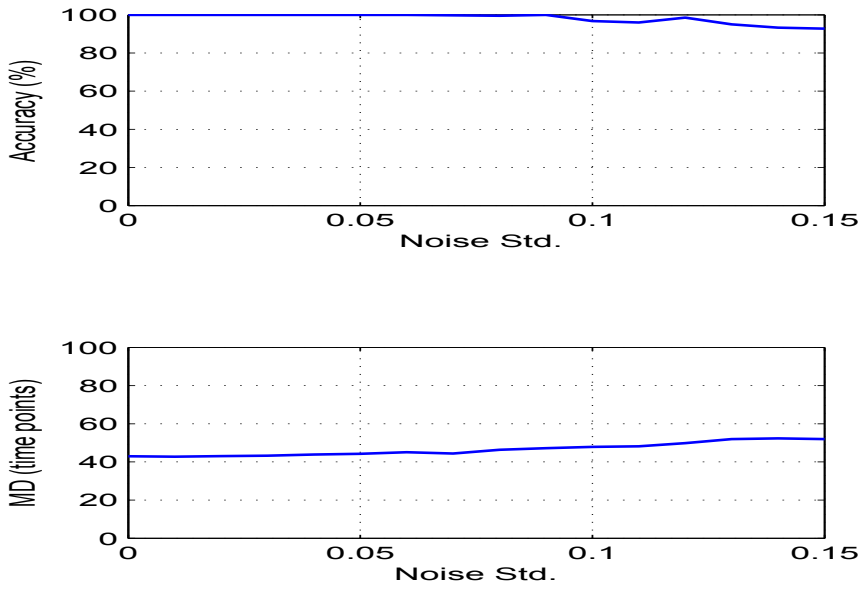

Figure 4.5: Robustness of the proposed method against different magnitudes of noise in the simulated data set. Noise Standard deviation (Noise Std.) vs. Overall Accuracy (top). Noise Standard deviation (Noise Std.) vs. Detection Delay (bottom). The unit of MD is time points = number of time points or number of observations.

all the methods. We note that all the methods can achieve accuracies close to $\kappa = 1$. The value of κ-coefficient increases with the increase in the threshold value, but as a consequence, the mean detection delay also increases as shown in Figure 4.3 (bottom). Figure 4.4 summarizes the plots of Figure 4.3 and gives a more obvious comparison by plotting MD against the corresponding κ-coefficients, for each method considered here. Figure 4.4 basically shows different trade offs between kappa (accuracy) and MD for each method, which is analogous to the Receiver's Operating Characteristics (ROC) curve that plots trade offs between TP and FP when only accuracies (without MD) are considered. Focusing on the significant region of this plot i.e. after $\kappa = 0.6$, we observe that although the difference between the plots is not very large, M1 and M4 perform slightly better than the rest of the two methods. This fact is also obvious from the comparison shown in Table 4.1 for the same accuracy of 99%.

We note that the results achieved here were according to our expectations because the simulated data lacks the effects of the complex natural phenomenon which are present in the real world NDVI data. Moreover, the noise in simulated data is Gaussian which satisfies the Gaussanity assumptions in M2-M4, and the points of actual changes in the training data are known exactly, hence no miss-labeling which can effect the training adversely. Nevertheless, these results illustrate the correctness of the approaches in that their tendency is to detect the real changes and avoid no change events.

The performance of M1 was also checked on simulated data sets with different magnitudes (standard deviations) of noise. The results are summarized in Figure 4.5. Figure 4.5 (top) shows that the accuracy drops very slightly with the increase in the standard deviation of the signal noise, and remains above 90% even at standard deviation as high as 0.15. Figure 4.5 (bottom) shows a very slight increase in the detection delay, from 42 at noise std. = 0 to 53 at noise std. = 0.15. These results show the robustness of the proposed framework M1 to different magnitudes of signal noise.

4.5.2 Results for Synthetic Data

We prepared 1000 change and 1000 no-change examples following the process explained in Section 4.4.2, and also used in [13, 17, 18, 50, 51]. A randomly selected set of 50% of the samples was taken as training set and the rest 50% as test set. The results have been summarized in Figure 4.6, Figure 4.7 and Table 4.2. Figure 4.6 (top) shows that the value of κ increases with the increase in the threshold, except for M4 that decreased after reaching its peak. The increase in threshold also causes increase in the detection delays, as shown in Figure 4.6 (bottom), because the test statistic has to attain bigger values in order to raise the change alarm. Figure 4.7 summarizes the plots of Figure 4.6 by plotting MD against the corresponding $\kappa-$coefficients, for each method considered here. The plots in Figure 4.7 are analogous to ROC curves and give a more obvious com-

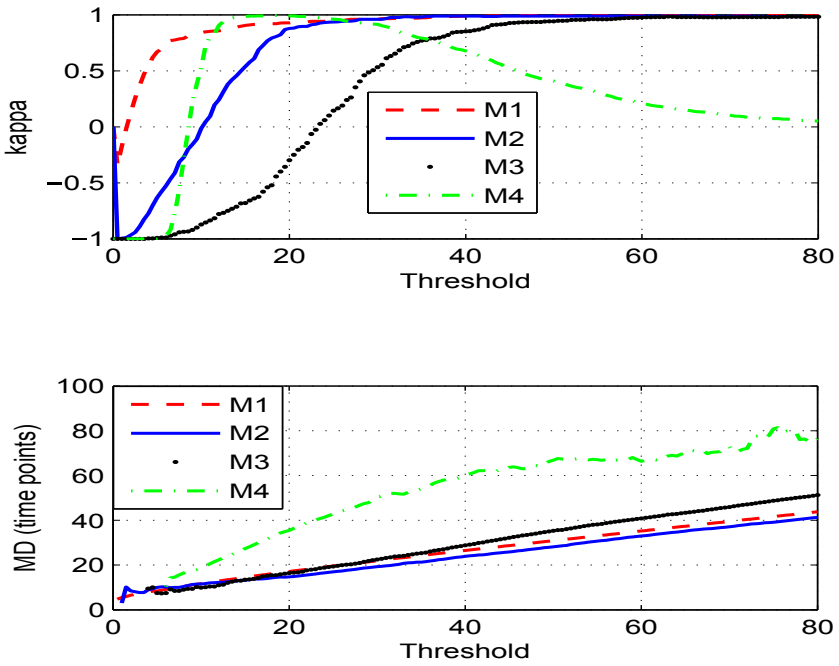

Figure 4.6: Comparison between performances of proposed (M1), original CUSUM (M2), near real-time disturbance detection (M3) and near real-time beetle infestation detection (M4) methods, on synthetic data. Threshold (λ) vs. κ-coefficient (top). Threshold (λ) vs. MD (bottom). The unit of MD is time points = number of time points or number of observations.

Table 4.2: Comparison between near real-time performances of our proposed method (M1) and the existing three methods, on synthetic data, at acceptable true positives, true negatives and accuracy.

Method	TP	TN	Acc.	MD	Threshold (λ)
M1	100	94	97	17	20.50
M2	97	97	97	17	25
M3	100	94	97	33	04.75
M4	100	94	97	23	02.50

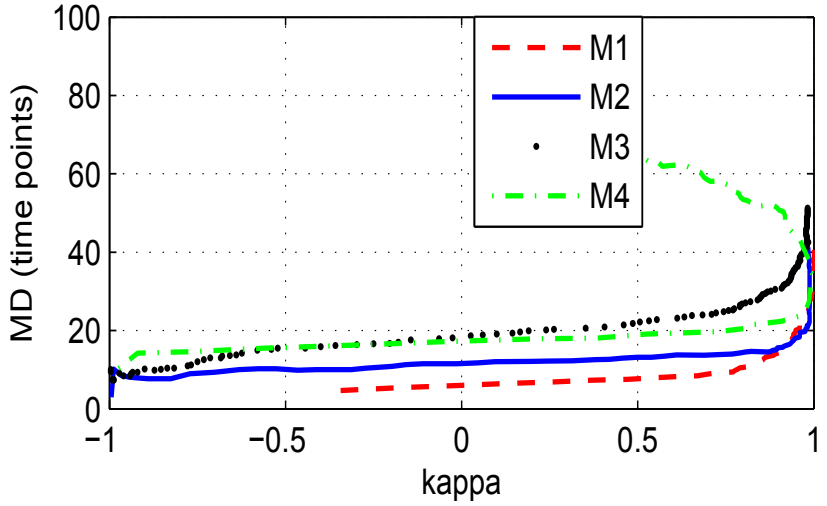

Figure 4.7: Comparison of "kappa-coefficient vs. MD" performances of proposed (M1), original CUSUM (M2), near real-time disturbance detection (M3) and near real-time beetle infestation detection (M4) methods on synthetic data. The unit of MD is time points = number of time points or number of observations.

parison. Here the difference between the performances is slightly more obvious than in case of simulated data because the data values and noise are real, only the changes introduced are synthetic i.e. due to concatenation of known no-change and change parts. Considering the significant part of the curves i.e. the region after $\kappa \approx 0.6$ on the horizontal axis, our proposed method M1 performs better than the rest until the point around $\kappa \approx 0.90$, where M1 and M2 become similar in performance. Table 4.2 also highlights this where detection delays are compared for all the four methods against the same accuracy of 97%. The M1 and M2 have similar results, better than M3 and M4. Furthermore, M4 performs better than M3. The reader should not be confused by the M4 curve hooking back in Figure (4.7). This behavior is quite possible because MD is not a function of the $\kappa-$coefficient. The $\kappa-$coefficient is calculated from TP, TN, FP and FN, which depend upon the threshold value. Two different thresholds can yield exactly the same $\kappa-$coefficient with different values of MD. The hooked curve shows exactly the same

Table 4.3: Comparison between near real-time performances of our proposed method (M1) and the existing three methods, on near real-time NDVI data, at acceptable true positives, true negatives and accuracy.

Method	TP	TN	Acc.	MD (tp)	Threshold (λ)
M1	100	90.6	95.3	79	26.1
M2	84	96	90	99	22
M3	90	90	90	131	5.8
M4	98	84	91	92	02

behavior i.e. same $\kappa-$coefficients with different values of MD.

The noise is far from Gaussian in this case, but still the change points in the training data are known exactly, hence no miss-labeling. Therefore, M2 still trains very well, hence small difference between the results of M1 and M2. The fact that M2 performs better than M3 and M4, can be attributed to the difference between the types of test statistics being used in these methods. The M2 uses CUSUM statistics which is more robust to non-Gaussanity and detects small/gradual changes earlier than the statistics based on simple likelihood ratios [36] as used in M3 and M4. The M4 performing better than M3 confirms the findings of [14]

4.5.3 Results for Real World MODIS NDVI (Beetle Infestation) Data

The simulated and synthetic data sets provided important insights but these data sets do not include all the complexities that are encountered in the real world data e.g. atmospheric variations, light variations, lack of information about the exact time points of the changes etc. Therefore, analysis on the real world data is necessary. We tested all the methods on 355 change and 355 no-change examples of beetle infestation data, collected as explained in Section 4.4.3.

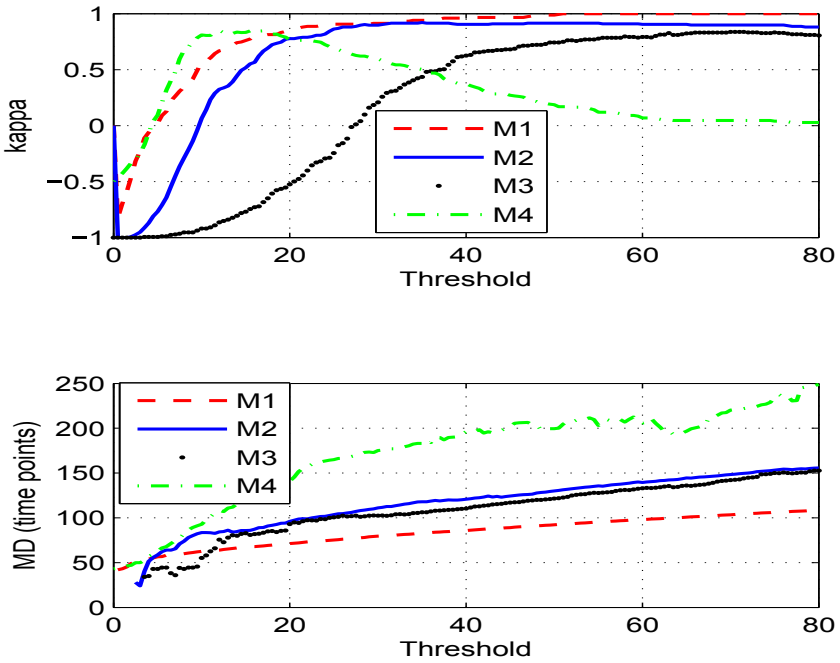

Figure 4.8: Comparison between performances of proposed (M1), original CUSUM (M2), near real-time disturbance detection (M3) and near real-time beetle infestation detection (M4) methods, on real-world data. Threshold (λ) vs. κ-coefficient (top). Threshold (λ) vs. MD (bottom). The unit of MD is time points = number of time points or number of observations.

The results have been summarized in Figure 4.8, Figure 4.9 and Table 4.3. As explained before, all the 4 methods have different ranges of threshold values, but we have scaled all of them to a single range of 0-80, for the sake of simplicity in comparison. The absolute values of the thresholds are not important here because we only want to graphically present the best possible performances by each of the methods considered here. Figure 4.8 shows similar trends for each method as in the case of synthetic data since the two data sets are close in nature to each other. Figure 4.8 (top) presents the "Threshold vs. $\kappa-$coefficient" plots of all the methods. The κ values of M2, M3 and M4 drop after reaching their peaks, whereas the κ value of M1 remains constant. Figure 4.8 (bottom) presents

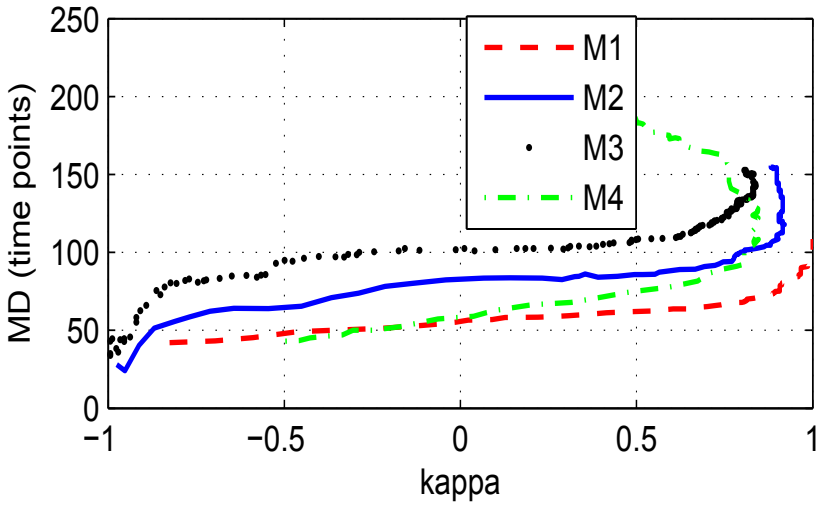

Figure 4.9: Comparison of "kappa-coefficient vs. MD" performances of proposed (M1), original CUSUM (M2), near real-time disturbance detection (M3) and near real-time beetle infestation detection (M4) methods on real-world data. The unit of MD is time points = number of time points or number of observations.

the "Threshold vs. MD" plots of all the methods. The MD generally increases with the increase in threshold value, for all the methods. Figure 4.9 summarizes the plots of Figure 4.8 by plotting κ-coefficient against the corresponding accuracies, for each method considered here. The plots in Figure 4.9 are analogous to ROC curves and give a clearer comparison. It can be seen in Figure 4.9 that our proposed method, M1, performs better than the other 3 methods by a significant margin. For the similar values of $\kappa-$coefficient, M1 incurs much lower detection delay than the rest of the 3 methods. Furthermore, M1 can still get close to $\kappa = 1$, unlike rest of the methods which peaked at significantly lower κ values than M1. Table 4.3 compares the performances of all the methods at acceptable accuracies (95.3% for M1, 90%, 90% and 91% for M2, M3 and M4, respectively). It can be noticed that M1 incurs much lower detection delay at a higher accuracy/kappa than rest of the three methods.

The reason behind such a significant difference between the performance of M1 and the

rest of the methods can be attributed to three facts: 1) The test statistics in M1 derived from the parameter time-series which has been shown to be better than the statistics derived from the raw NDVI time-series [14]. 2) RSPRT statistics are used in M1 which are more robust to non-Gaussanity and detects small changes faster than the simple statistics based on likelihood ratios [36]. 3) The likelihood ratios used in deriving RSPRT statistics were estimated directly using RULSIF algorithm [27], which performs better than the likelihood ratios derived from individual density functions which are based on Gaussanity assumption or estimated individually using Gaussian kernels. Furthermore, it is very difficult to obtain the exact ground reference data for long time-series [56]. The forestry departments which are monitoring those forests also confirmed that the survey maps were manual and subjected to errors, hence the real world data did not have exact information about the change points in each time-series. Therefore, we took $t = 230$ as common reference point for all the methods i.e. the last point for known no-change part in every time-series, from which MDs were calculated. However, this does not mean that the changes were known to have occurred at that point. Majority of the time-series changed at later unknown points which implies that the training data had miss-labeling at some points after $t = 230$. This suggests that M1 is more robust to miss-labeling in the training data as compared to the other 3 methods.

4.5.4 Cross Validation and Automatic Threshold Tuning Results

We performed cross validation experiments of the proposed method M1 on all the three data sets. The cross validation consisted of 10 runs, where in each run 50% of the data samples selected randomly was used as training and the rest 50% as test sets. The results have been summarized in Table 4.4. The table consists of 2 vertical halves and 3 horizontal parts. The left vertical half summarizes the mean of performance indices for the

Table 4.4: Results of 50% cross-validations (10 runs) of M1 on all the three data sets. In every run randomly selected 50% of the data samples of a particular data set is taken as training set, and the rest as test set. TP = True Positive, TN = True Negative, Acc. = Overall Accuracy, κ= kappa-coefficient, MD = Mean Detection Delay, λ = threshold value, Sim. = Simulated data set, Syn. = Synthetic data set and R.W. = Real-World NDVI beetle infestation data. The units of TP, TN and Acc. are "%", and that of MD is tp = number of time points or observations.

Dataset	Data	Mean (10 runs)						Standard Deviation (10 runs)					
		TP	**TN**	**Acc.**	κ	**MD (tp)**	λ	**TP**	**TN**	**Acc.**	κ	**MD (tp)**	λ
Sim.	Training	100	95.75	97.9	0.957	45.7	16.67	0	1.09	0.54	0.011	0.28	0.52
	Test	100	96.00	98.00	0.960	45.8		0	1.26	0.63	0.013	0.63	
Syn.	Training	100	95.5	97.75	0.955	21.8	29.83	0	1.47	0.74	0.015	0.25	0.77
	Test	100	95.2	97.30	0.952	22.7		0	1.34	0.68	0.013	0.51	
R.W.	Training	100	89.50	94.75	0.895	80.5	27.42	0	2.4	1.2	0.024	0.47	2.1
	Test	100	88.00	94.00	0.880	79.1		0	2.5	1.25	0.025	0.7	

training and test sets of all the data sets, whereas the right vertical half summarizes the standard deviations of the 10 runs. Each of the three horizontal parts of the table summarizes the performance indicators for one of the three data sets. The value of ψ was selected, in the first run of the cross validation experiment on simulated data, against 98% accuracy and kept fixed for rest of the experiments. The mean performance indices of the training set and test set are very close to each other, in case of all the three data sets. This suggests that the thresholds selected from the training data sets are robust and perform equally well on the unseen test data.

The threshold selected in each run of this cross validation experiment was selected automatically by our proposed threshold tuning technique (based on $\kappa-$coefficient and MD) as explained in Section 4.2.2. Apart from tuning, another advantage of this technique, especially in cross-validation experiment, is that it can be incorporated in the code and a complete set of cross-validation results can be generated in a single go, without stopping and selecting the right threshold in each cross-validation run. The table shows that the thresholds selected by this technique on the training data sets, perform nearly similar on

the test data sets. This shows the effectiveness of the proposed threshold tuning technique. It is worth noting that the range of the suitable threshold values may change from one data set to another, hence needs tuning on all the data sets separately. However, the proposed threshold tuning technique avoids this problem since the value of ψ selected for one data set holds good for rest of the data sets as well.

4.6 Conclusion

In this paper, we proposed a supervised framework for near real-time land cover change detection which uses EKF to fit a triply modulated cosine function to a MODIS NDVI time-series, extracts its time varying parameters and derives the RSPRT test statistics from the trend parameter. Instead of using traditional likelihood ratios, we exploited the usefulness of relative density ratios estimated directly using RULSIF algorithm as proposed in [27], in deriving the RSPRT statistics. Our framework slightly reduces the correlation in the parameter time series, and, unlike CUSUM formulation in [13, 41], deals with the no-change samples as identically distributed which is an important assumption of CUSUM. We tested the framework on three different data sets, against different noise level, and also performed cross-validation. Furthermore, we compared its performance with 3 recently published near real-time change detection methods in remote sensing literature.

Our analysis of the proposed method on different data sets considered here, and also its comparison with 3 published methods, helped us in finding the answers to the questions we raised earlier in this manuscript. Our findings can be summarized as: I) The promising results of the proposed method (M1) suggest that direct estimation of relative density ratios, from the data, is a viable option for supervised classification of remote sensing time series data, II) The results of comparison between M1 and M2 [13] suggest that RSPRT/CUSUM statistics, when derived from the parameter time series instead of the

raw data, achieve significant improvement in the performance, III) Utilizing the benefits of parameter time-series [14, 16], RSPRT statistics and relative density ratio estimation simultaneously, enabled the framework to incur lower detection delays, with higher accuracy than the rest of the methods (M2 - M4) which use traditional likelihood ratios with individual densities either estimated using Gaussian kernels [13] or assumed to be Gaussian in nature [10, 14]. The difference in the results was small in case of simulated and synthetic data sets because some assumptions involved in the derivation of M2-M4 were satisfied to some extent e.g. Gaussanity assumption, and there was no miss-labeling. However, the difference became more significant in case of the real world beetle infestation data when the Gaussanity assumption was violated and there was considerable miss-labeling as well in the training data. This also suggests that the proposed framework is more robust to miss-labeling as compared to the other 3 the methods.

We also proposed a simple heuristic technique for automatic threshold tuning in near real-time change detection framework. Unlike commonly considered 2 indices (FP and FN), this technique considers 3 performance indices (FP, FN and MD), which are challenging to deal with simultaneously while tuning the threshold. This technique proved useful in cross-validation experiments and allowed us to generate the whole set of results in a single execution, without having to select the thresholds manually in each run. The threshold values presented in Table 4.4 were selected automatically by the framework using this technique, which verifies that it can tune the thresholds successfully.

Bibliography

[1] G. W. Meigs, R. E. Kennedy, and W. B. Cohen, "A landsat time series approach to characterize bark beetle and defoliator impacts on tree mortality and surface fuels in conifer forests," *Remote Sensing of Environment*, vol. 115, pp. 3707–3718, Dec 2011.

[2] A. J. Meddens, J. A. Hicke, L. A. Vierling, and A. T. Hudak, "Evaluating methods to detect bark beetle-caused tree mortality using single-date and multi-date landsat imagery," *Remote Sensing of Environment*, vol. 132, pp. 49–58, May 2013.

[3] A. J. Meddens, J. A. Hicke, and L. A. Vierling, "Evaluating the potential of multispectral imagery to map multiple stages of tree mortality," *Remote Sensing of Environment*, vol. 115, pp. 1632–1642, Jul 2011.

[4] N. R. Goodwin, N. C. Coops, M. A. Wulder, S. Gillanders, T. A. Schroeder, and T. Nelson, "Estimation of insect infestation dynamics using a temporal sequence of landsat data," *Remote Sensing of Environment*, vol. 112, pp. 3680–3689, Sep 2008.

[5] N. R. Goodwin, S. Magnussen, N. C. Coops, and M. A. Wulder, "Curve fitting of time-series landsat imagery for characterizing a mountain pine beetle infestation," *International Journal of Remote Sensing*, vol. 31, pp. 3263–3271, Jun 2010.

[6] N. C. Coops, S. N. Gillanders, M. A. Wulder, S. E. Gergel, T. Nelson, and N. R. Goodwin, "Assessing changes in forest fragmentation following infestation using

time series landsat imagery," *Forest Ecology and Management*, vol. 259, pp. 2355–2365, May 2010.

[7] N. C. Coops, M. Johnson, M. A. Wulder, and J. C. White, "Assessment of quickbird high spatial resolution imagery to detect red attack damage due to mountain pine beetle infestation," *Remote Sensing of Environment*, vol. 103, pp. 67–80, Jul 2006.

[8] M. A. Wulder, J. C. White, N. C. Coops, and C. R. Butson, "Multi-temporal analysis of high spatial resolution imagery for disturbance monitoring," *Remote Sensing of Environment*, vol. 112, pp. 2729–2740, Jun 2008.

[9] J. White, W. M.A., and D. Grills, "Detecting and mapping mountain pine beetle red-attack damage with spot-5 10-m multispectral imagery," *BC Journal of Ecosystems and Management*, vol. 7, no. 2, pp. 105 – 118, 2006.

[10] J. Verbesselt, A. Zeileis, and M. Herold, "Near real-time disturbance detection using satellite image time series," *Remote Sensing of Environment*, vol. 123, pp. 98–108, Aug 2012.

[11] J. Verbesselt, R. Hyndman, G. Newnham, and D. Culvenor, "Detecting trend and seasonal changes in satellite image time series," *Remote Sensing of Environment*, vol. 114, pp. 106–115, Jan 2010.

[12] J. Verbesselt, R. Hyndman, A. Zeileis, and D. Culvenor, "Phenological change detection while accounting for abrupt and gradual trends in satellite image time series," *Remote Sensing of Environment*, vol. 114, pp. 2970–2980, Dec 2010.

[13] T. L. Grobler, E. R. Ackermann, A. J. Van Zyl, J. C. Olivier, W. Kleynhans, and B. P. Salmon, "Using page's cumulative sum test on modis time series to detect land-cover changes," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, pp. 332–336, March 2013.

[14] A. Anees and J. Aryal, "Near-real time detection of beetle infestation in pine forests

using modis data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 99, pp. 1–11, 2014.

[15] A. Anees, J. Olivier, M. O'Rielly, and J. Aryal, "Detecting beetle infestations in pine forests using modis ndvi time-series data," in *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International*, pp. 3329–3332, July 2013.

[16] A. Anees and J. Aryal, "A statistical framework for near-real time detection of beetle infestation in pine forests using modis data," *Geoscience and Remote Sensing Letters, IEEE*, vol. 11, pp. 1717–1721, Oct 2014.

[17] W. Kleynhans, J. C. Olivier, K. J. Wessels, B. P. Salmon, F. van den Bergh, and K. Steenkamp, "Detecting land cover change using an extended kalman filter on modis ndvi time-series data," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, pp. 507–511, May 2011.

[18] W. Kleynhans, J. C. Olivier, K. J. Wessels, F. van den Bergh, B. P. Salmon, and K. C. Steenkamp, "Improving land cover class separation using an extended kalman filter on modis ndvi time-series data," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, pp. 381–385, Apr 2010.

[19] R. Watson, I. Noble, B. Bolin, N. Ravindranath, D. Verardo, and D. Dokken, "Land Use, Land-Use Change and Forestry," *Cambridge University Press, UK*, p. 375, 2000.

[20] L. Eklundh, T. Johansson, and S. Solberg, "Mapping insect defoliation in scots pine with modis time-series data," *Remote Sensing of Environment*, vol. 113, pp. 1566–1573, Jul 2009.

[21] Y. Fang, A. R. Ganguly, N. Singh, V. Vijayaraj, N. Feierabend, and D. T. Potere, "Online change detection: Monitoring land cover from remotely sensed data," in

Data Mining Workshops, 2006. ICDM Workshops 2006. Sixth IEEE International Conference on, pp. 626–631, IEEE, 2006.

[22] Y. Kang, "Real-time change detection in time series based on growing feature quantization," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–6, June 2012.

[23] Y. Kawahara and M. Sugiyama, "Change-Point Detection in Time-Series Data by Direct Density-Ratio Estimation," in *Proceedings of the SIAM International Conference on Data Mining, SDM 2009, April 30 - May 2, 2009, Sparks, Nevada, USA*, pp. 389–400, SIAM, 2009.

[24] S. Liu, M. Yamada, N. Collier, and M. Sugiyama, "Change-point detection in time-series data by relative density-ratio estimation," *Neural Networks*, vol. 43, no. 0, pp. 72 – 83, 2013.

[25] M. Sugiyama, T. Suzuki, S. Nakajima, H. Kashima, P. von Bünau, and M. Kawanabe, "Direct importance estimation for covariate shift adaptation," *Annals of the Institute of Statistical Mathematics*, vol. 60, no. 4, pp. 699–746, 2008.

[26] T. Kanamori, S. Hido, and M. Sugiyama, "A least-squares approach to direct importance estimation," *J. Mach. Learn. Res.*, vol. 10, pp. 1391–1445, Dec. 2009.

[27] M. Yamada, T. Suzuki, T. Kanamori, H. Hachiya, and M. Sugiyama, "Relative density-ratio estimation for robust distribution comparison," *Neural computation*, vol. 25, no. 5, pp. 1324–1370, 2013.

[28] A. Gretton, A. Smola, J. Huang, M. Schmittfull, K. Borgwardt, and B. Schölkopf, *Dataset Shift in Machine Learning*, ch. 8, pp. 131–160. Cambridge, MA, USA: MIT Press, 2009.

[29] S. Bickel, M. Brückner, and T. Scheffer, "Discriminative learning for differing train-

ing and test distributions," in *Proceedings of the 24th International Conference on Machine Learning*, ICML '07, (New York, NY, USA), pp. 81–88, ACM, 2007.

[30] T. Kanamori, T. Suzuki, and M. Sugiyama, "Statistical analysis of kernel-based least-squares density-ratio estimation," *Machine Learning*, vol. 86, no. 3, pp. 335–367, 2012.

[31] T. Kanamori, T. Suzuki, and M. Sugiyama, "Computational complexity of kernel-based density-ratio estimation: A condition number analysis," *Mach. Learn.*, vol. 90, pp. 431–460, Mar. 2013.

[32] M. Sugiyama, T. Suzuki, and T. Kanamori, "Density-ratio matching under the bregman divergence: a unified framework of density-ratio estimation," *Annals of the Institute of Statistical Mathematics*, vol. 64, no. 5, pp. 1009–1044, 2012.

[33] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[34] J. Im, J. Rhee, J. R. Jensen, and M. E. Hodgson, "An automated binary change detection model using a calibration approach," *Remote Sensing of Environment*, vol. 106, no. 1, pp. 89 – 105, 2007.

[35] J. Cohen, "A Coefficient of Agreement for Nominal Scales," *Educational and Psychological Measurement.*, vol. 20, no. 1, p. 37, 1960.

[36] D. Montgomery, *Introduction to statistical quality control*. New York, USA: Wiley, 3rd ed., 1997.

[37] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1/2, pp. pp. 100–115, 1954.

[38] R. S. Lunetta, J. F. Knight, J. Ediriwickrema, J. G. Lyon, and L. D. Worthy, "Land-cover change detection using multi-temporal modis ndvi data," *Remote Sens. Environ.*, vol. 105, pp. 142–154, Nov 2006.

[39] W. A. Shewhart and W. E. Deming, *Statistical method from the viewpoint of quality control*. New York : Dover, 1986.

[40] M. Sugiyama, S. Nakajima, H. Kashima, P. V. Buenau, and M. Kawanabe, "Direct importance estimation with model selection and its application to covariate shift adaptation," in *Advances in neural information processing systems*, pp. 1433–1440, 2008.

[41] T. L. Grobler, *Sequential and non-sequential hypertemporal classification and change detection of Modis time-series*. PhD thesis, University of Pretoria, South Africa., 2012.

[42] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sensing of Environment*, vol. 37, no. 1, pp. 35 – 46, 1991.

[43] A. Stein, J. Aryal, and G. Gort, "Use of the bradley-terry model to quantify association in remotely sensed images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, pp. 852–856, April 2005.

[44] R. G. Pontius Jr and M. Millones, "Death to kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment," *International Journal of Remote Sensing*, vol. 32, no. 15, pp. 4407–4429, 2011.

[45] A. Agresti, *An Introduction to Categorical Data Analysis*. Wiley Series in Probability and Statistics, Wiley, 1996.

[46] A. Zeileis, F. Leisch, C. Kleiber, and K. Hornik, "Monitoring structural change in dynamic econometric models," *Journal of Applied Econometrics*, vol. 20, no. 1, pp. 99–121, 2005.

[47] A. Zeileis, A. Shah, and I. Patnaik, "Testing, monitoring, and dating structural changes in exchange rate regimes," *Comput. Stat. Data Anal.*, vol. 54, pp. 1696–1706, June 2010.

[48] F. Leisch, K. Hornik, and T. U. Wien, "Monitoring structural changes with the generalized fluctuation test. econometric theory," *Econometric Theory*, vol. 16, pp. 835 – 854, 2000.

[49] T. Grobler, E. Ackermann, A. van Zyl, J. Olivier, W. Kleynhans, and B. Salmon, "An inductive approach to simulating multispectral modis surface reflectance time series," *Geoscience and Remote Sensing Letters, IEEE*, vol. 10, pp. 446–450, May 2013.

[50] B. P. Salmon, J. C. Olivier, K. J. Wessels, W. Kleynhans, F. van den Bergh, and K. C. Steenkamp, "Unsupervised land cover change detection: Meaningful sequential time series analysis," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, pp. 327–335, Jun 2011.

[51] B. Salmon, J. Olivier, W. Kleynhans, K. Wessels, F. van den Bergh, and K. Steenkamp, "The use of a multilayer perceptron for detecting new human settlements from a time series of modis images," *International Journal of Applied Earth Observation and Geoinformation*, vol. 13, pp. 873–883, Dec 2011.

[52] "http://www.for.gov.bc.ca/ftp/hre/external/publish/web/bcmpb/year9/bcmpb.v9.2011kill.pdf."

[53] "http://www.for.gov.bc.ca/hfp/health/overview/mpb_h-istory.htm."

[54] "http://foresthealth.fs.usda.gov/portal/flex/ids."

[55] "http://landweb.nascom.nasa.gov/cgi-bin/developer/tilemap.cgi."

[56] G. Jianya, S. Haigang, M. Guoruo, and Z. Qiming, "A review of multi-temporal remote sensing data change detection algorithms," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37.B7, pp. 757 – 762, 2008.

5 A Robust Multi-Kernel Change Detection Framework for Detecting Leaf Beetle Defoliation using Landsat 7 ETM+ Data[1]

Overview

A robust non-parametric framework, based on multiple Radial Basic Function (RBF) kernels, is proposed in this study, for detecting land/forest cover changes using Landsat 7 ETM+ images. One of the widely used frameworks is to find change vectors (difference image) and use a supervised classifier to differentiate between change and no-change. The Bayesian Classifiers e.g. Maximum Likelihood Classifier (MLC), Naive Bayes (NB), are widely used probabilistic classifiers which assume parametric models, e.g. Gaussian function, for the class conditional distributions. However, their performance can be limited if the data set deviates from the assumed model. The proposed framework exploits the useful properties of Least Squares Probabilistic Classifier (LSPC) formulation i.e. non-parametric and probabilistic nature, to model class posterior probabilities of the difference image using a linear combination of a large number of Gaussian kernels. To this

[1]This chapter has been submitted as a journal article in ISPRS Journal of Photogrammetry and Remote Sensing.

end, a simple technique, based on 10-fold cross-validation is also proposed for tuning model parameters automatically instead of selecting a (possibly) suboptimal combination from pre-specified lists of values. The proposed framework has been tested and compared with Support Vector Machine (SVM) and NB for detection of defoliation, caused by leaf beetles (*Paropsisterna* spp) in *Eucalyptus nitens* and *E. globulus* plantations of two test areas, in Tasmania, Australia, using raw bands and band combination indices of Landsat 7 ETM+. It was observed that due to multi-kernel non-parametric formulation and probabilistic nature, the LSPC outperforms parametric NB with Gaussian assumption in change detection framework, with Overall Accuracy (OA) ranging from 93.6% ($\kappa = 0.87$) to 97.4% ($\kappa = 0.94$) against 85.3% ($\kappa = 0.69$) to 93.4% ($\kappa = 0.85$), and is more robust to changing data distributions. Its performance was comparable to SVM, with added advantages of being probabilistic and capable of handling multi-class problems naturally with its original formulation.

5.1 Introduction

Land cover change detection using satellite imagery has been one of the important areas of research over the past few decade [1–3]. One of the main focuses of the this research has been forest cover changes induced by insect attacks/infestations [4–9] because they have both environmental and economic implications e.g. increased risk of forest fires, less carbon absorption, loss of valuable wood etc [10, 11]. Monitoring large forests by manual surveying can be very time consuming, expensive and labor intensive, sometimes impossible. This necessitates development of cost effective, efficient and quick automated alarm systems for the concerned authorities in order to take timely actions to mitigate the insect infestations in the affected parts of the forests [4, 6]. Therefore, numerous studies applying satellite imagery to detect vegetation cover changes due to different phenomena e.g. forest clear cutting, fire and insect attacks, have been proposed in the literature

[4, 7–9, 12, 13]. These studies showed that, due to availability of a variety of sensors and advancements in the computing technologies, remote sensing is a viable option for tackling this problem [7]. An important consideration, before developing any satellite imagery based change detection algorithm, is the choice of the satellite imagery [2, 3]. Different satellite images have different spatial, temporal and spectral resolutions. Therefore, the choice of satellite imagery is strictly application dependent e.g. fine spatial resolution data suit land cover changes at local scale which may be too small to be captured by moderate or coarse spatial resolution images. Similarly, moderate resolution satellite images suit regional land cover dynamics, and coarse spatial resolution imagery can be helpful in global scale change detection [1].

In this paper, the main focus is on usage of Landsat 7 ETM+ (Enhanced Thematic Mapper Plus) 30m resolution imagery for detecting defoliation (change) in *Eucalyptus nitens* and *E. globulus* plantations of north-eastern Tasmania, Australia due to leaf beetles (*Paropsisterna* spp). Existing EO literature contains considerable number of algorithms proposed for detecting changes in vegetation cover using multi-temporal satellite images, such as visual inspection [13, 14], image differencing [12, 15], image rationing [16], change vector analysis (CVA) [17], image regression [18], multi-temporal linear data transformation [15, 19, 20] e.g. Principle Component Analysis (PCA) and Tasseled Cap Transformation (TCT). Performance of these methods depend on the characteristics of the study areas and images used in the analysis [3]. Therefore, despite the existence of a variety of alternatives with their respective pros and cons, all of them are application specific and no single change detection algorithm can be declared as optimal for all the cases [3].

Image differencing and CVA have been the most widely used change detection algorithms [2, 3]. A major difficulty of image differencing or CVA based methods is to tune an optimal threshold [2, 3, 21]. Another issue with such methods is that they are simple and may not separate real changes from many other difference between the two-date imagery

due to various factors [4]. These misleading changes may be due to non-normalized radiometric differences, non-linear noise, changes due to atmospheric conditions [22] and changes caused by vegetation phenology, growth and senescence etc. [5]. Such changes are not linearly separable from the real changes in spectral space and may require more sophisticated techniques for processing in higher dimensional spaces [22]. Therefore, many authors have developed different algorithms which use image differencing or CVA as basic step and classify the difference image or Change Vectors (CVs) into change and no-change efficiently [21]. Among these, the algorithms which use supervised classifiers e.g. artificial neural networks (ANN) [23, 24], support vector machines (SVM) [25, 26], and Bayesian classifiers e.g. Maximum Likelihood Classifier (MLC) and Naive Bayes (NB) [27–31] etc., have gained significant importance over the recent past.

Bayesian Classifiers [28, 30, 31] well known and commonly used probabilistic classifiers in EO community for change detection and classification [27, 29, 32, 33]. They are capable of multi-class classification and assigning probabilistic class membership to the test samples. However, their major drawback is that they rely on a parametric density model (mostly Gaussian) for the data in each class, which may be misleading and far from being true in practical cases [34]. As a consequence, it performs very well as long as the data is following the assumed parametric model, but becomes sub-optimal when the assumption is violated significantly [33, 35, 36]. Therefore, a change detection algorithm, that is more robust in the cases where the data is more likely unknown and far from assumed parametric distribution (most of the practical scenarios), will prove useful not only for vegetation data but also for broader range of applications on EO data.

There exist many classifiers in the literature, e.g. Neural Networks (NN) [37–40] and Support Vector Machines (SVM) [36, 37, 41–44], which are non-parametric. SVM is the most popular of them in the EO community [34]. It uses kernel trick to transform features into high-dimensional Hilbert space and achieves a linear decision boundary

in high-dimensional space, for the data which may not be linearly separable in spectral space. Many researchers have proposed other kernel based non-parametric change detection approaches, some of which use a one-class version of SVM. Volpi *et al.* [45] used kernel k-means on difference image and kernel image differencing for unsupervised change detection in satellite images. Shah-Hosseini *et al.* [22, 46] used kernels to transform images from spectral to feature space, and used kernel clustering and Support Vector Domain Description (SVDD) on initial pseudo training set to achieve unsupervised change detection maps. Bovolo *et al.* used a combination of kernels with SVDD [47] and Semi-Supervised Support Vector Machine (S^3VM) [48] on difference image, with Bayesian initial thresholding to obtain pseudo training set for unsupervised change detection. All of these studies showed promise of kernel based methods in remote sensing change detection. The advantages of these methods are that they are either unsupervised or semi-supervised. However, such methods and the classical SVM only learn the decision boundary and cannot provide class membership confidence (probabilities) to a test sample [49, 50]. Moreover, the original formulation of the provides only binary classification, hence different strategies need to be implemented in order to use it for multi-class problems. Therefore, a classification technique that is not only non-parametric, but also capable of handling multi-class problems and providing probabilistic output (assigning degree of class-membership to test samples), may prove very useful in broad range of remote sensing applications, if it performs comparably to SVM.

The LSPC [49, 50] is a relatively new classifier introduced recently. It is non-parametric like SVM, with the added advantages of multi-class handling and probabilistic nature. It models class posterior probabilities using a linear combination of a large number of weighted Gaussian Kernels, which is more flexible and could provide a better approximation of the distributions that may be far from the assumed parametric models normally used in traditional Bayesian classifiers e.g. Gaussian distribution [34]. Although, the

computational efficiency has not been crucial in this study, it is worth mentioning here that it has been shown in [49, 50] that after certain simplifications, the LSPC is computationally more efficient than SVM for multi-class problems with large data sets, because of its ability to obtain its solution in a class-wise manner by solving a regularized system of linear equations analytically. In a broader sense it can be seen as a non-parametric technique for estimating the conditional distribution, and can be used in numerous other cases where normally assumptions are made about the conditional distributions of the data. However, its usefulness has not been exploited in remote sensing applications yet. All these facts make LSPC an interesting concept to be introduced to the EO community. In this study, the aim is to: 1) devise a robust non-parametric change detection framework, by exploiting useful properties of the LSPC, for detect leaf beetle infestation in *Eucalyptus nitens* and *E. globulus* plantations using Landsat 7 ETM+ data, 2) analyze the potential and robustness of LSPC [49, 50], in non-parametric modeling of class posterior probabilities of EO data sets with unknown distributions, and 3) design an automated model selection procedure based on 10-fold cross validation for near-optimal tuning of the kernel parameters and the decision threshold.

The proposed algorithm first calculates CVs from two Landsat images (raw bands or band combination indices) and then applies the supervised LSPC formulation [49, 50] for differentiating between change and no-change classes. The weights of the kernels are learnt from the training data conveniently by evaluating a closed form expression derived from a system of linear equations [49, 50]. The parameters of Gaussian kernels can be derived using 10-fold cross validation experiment on the training data set[49, 50]. These parameters are normally selected as the best combination from pre-defined lists of values, hence leaving room for improvement since the best values may not be available in the lists. To this end, a simple optimization problem is devised for model selection, that can be solved using any constrained optimization algorithm to obtain the kernel parameters and thresh-

old. The advantage of such model selection is that it avoids trial and error procedure of model selection, which normally involves manual selection of lists of parameters values which may be suboptimal. The performance of the proposed algorithm is tested and compared with NB and SVM in detecting changes in Landsat raw bands and band combination indices data sets of two different test areas of north-eastern Tasmania, Australia, where *Eucalyptus nitens* and *E. globulus plantations* suffered with defoliation due to leaf beetles (*Paropsisterna* spp). The purpose of using raw bands and band combination indices data sets was to evaluate the robustness of both the methods in case of significantly different distributions of the data. This framework is an extension of a similar approach [6], that used multiple kernels for estimating the relative density ratios [51–54] directly from the EO satellite time series data and showed promise in land cover change detection problem.

5.2 Data and Methods

5.2.1 Data Description

The study area consists of *Eucalyptus Nitens* and *E. Globulus* plantations located in north-eastern Tasmania (147°15′0″E to 148°15′0″E and 41°0′0″S to 41°40′0″S, see Figure 5.1). The shape files containing polygons indicating damaged areas were provided by the field surveyors. These shape files were prepared by dedicated professionals after close and careful regular monitoring of the *Eucalyptus Nitens* and *E. Globulus* stands managed and looked after by Forestry Tasmania. Two cloud free Landsat 7 ETM+ images dated 17th May, 2009 and 20th May, 2010 (before and after the mentioned date of damage) were acquired from NASA's website. The Landsat images contained black strips, caused by scan line/sensor failure. Due to availability of sufficient ground truth examples in valid regions, these scan line errors were masked out and the pixel locations in these strips were not included in the analysis, in order to make the results reliable. The procedure of

preparing the masks for scan line errors, can be found on official website of United States Geological Survey (USGS).

First the Digital Numbers (DNs) were converted into radiance by the following formula [55, 56]

$$L_b = (G_b \times DN) + \xi_b, \tag{5.1}$$

where L_b is the radiance of band b, G_b is the gain and ξ_b is the bias. The radiance values were then converted in to Top of Atmosphere (TOA) reflectance using [56, 57]

$$R_b = \frac{\pi L_b d^2}{E_{sun,b} \sin \vartheta}, \tag{5.2}$$

where R_b is the TOA reflectance for band b, d is the earth-sun distance in astronomical units, $E_{sun,b}$ is the mean exoatmospheric solar irridiance at the specific band and ϑ is solar zenith angle. The above radiometric calibrations were carried out using ENVI 5.2 software. After calculating TOA reflectance and co-registration, the ArcGis software was used to prepare point data with 30m grid, from the provided polygons in the shape files. This point data was then overlayed on the corrected Landsat 7 ETM+ images at the exact locations. Two separate areas containing the marked damages were selected and their ground truth masks were prepared. These areas were treated as two separate data sets, namely, Area 1 and Area 2 containing 8764 and 8326 infested/defoliated pixels, respectively. Half of the available ground truth was used as training set and the rest as test set, for both the areas. These areas were further subdivided into two data sets, namely, 6 raw bands (bands 1-5 and band 7) and 7 band combination indices namely Land Surface Water Index (LSWI) [58], Normalized Difference Vegetation Index (NDVI) [59, 60], Modified Normalized Difference Wetness Index (MNDWI) [61], Enhanced Vegetation Index (EVI) [62], TCT-brightness, TCT-greenness and TCT-wetness [63]. These band combination indices were selected on the basis of their relevance to vegetation dynamics while ensuring

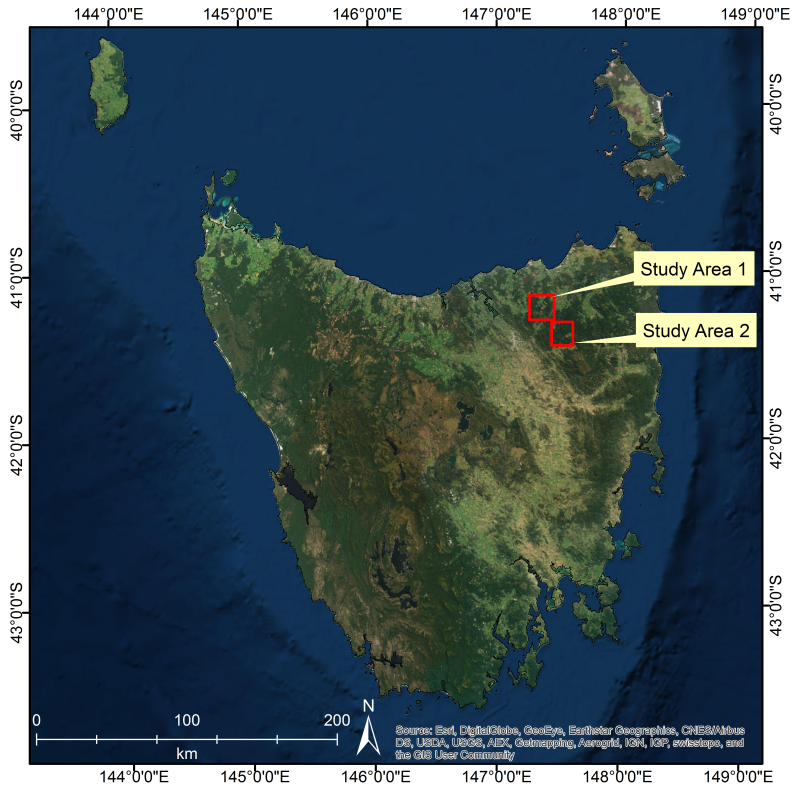

Figure 5.1: Map of Tasmania, Australia. The small boxes show the study areas located in North-East of Tasmania.

inclusion of each thematic band at least once.

The purpose of making these two data sets was to test the robustness of both the probabilistic classifiers on the data sets that are distributed significantly differently. Some previous studies have shown that Gaussian assumption has given reliable results on raw bands of Landsat images e.g. [27], hence band indices data set would be a good tester for robustness.

5.2.2 Least Squares Probabilistic Classifier for EO change detection in Landsat 7 ETM+ Images

Let $\mathbf{I}_1, \mathbf{I}_2 \left(\subset \mathbb{R}^{M \times N \times B}\right)$ be two images (before and after the change/defoliation event) with M rows, N columns and B bands/features. First each band/feature image $\mathbf{I}_t^b$ ($b \in$

$\{1,2,\ldots,B\}$, $t \in \{1,2\}$) is normalized and the difference image is calculated as [1]

$$\triangle \mathbf{I}^b = \mathbf{I}_2^b - \mathbf{I}_1^b. \tag{5.3}$$

Each vector at location (m,n) of the difference image $\triangle \mathbf{I}$,

$$\triangle \mathbf{I}(m,n) = \left[\triangle \mathbf{I}^1(m,n), \triangle \mathbf{I}^2(m,n), \ldots, \triangle \mathbf{I}^B(m,n)\right]^{\mathrm{T}}, \tag{5.4}$$

where $m \in \{1,2,\ldots,M\}$ and $n \in \{1,2,\ldots,N\}$, can be considered as an input feature vector denoted by $\mathbf{x} \in \mathscr{F}$ $\left(\subset \mathbb{R}^B\right)$ and its class label given by $c \in \mathscr{Y} = \{1,2,\ldots,C\}$, where C is the number of classes. Each input-output pair $(\mathbf{x},c)$ belongs to a joint probability distribution on $\mathscr{F} \times \mathscr{Y}$ with joint probability density $p(\mathbf{x},c)$. The aim here is to estimate the class posterior probability $p(c \mid \mathbf{x})$ from the training set $\mathbf{X}_{tr}$, containing training input-output pairs, given as

$$\mathbf{X}_{tr} = \{(\mathbf{x}_i, c_i) \in \mathscr{F} \times \mathscr{Y}\}_{i=1}^{N_{tr}}, \tag{5.5}$$

where N_{tr} is the number of samples in the training set $\mathbf{X}_{tr}$.

Popular probabilistic classifiers in remote sensing, Bayesian classifiers, rely on an assumption of a parametric function to model the class conditional distribution [28–31], and maximizes the log-likelihood over the data set to obtain the model parameters. Since the distribution of EO data is often unknown and can deviate significantly from the assumed parametric model causing poor classification accuracy [33–35], it is proposed here that the class posterior probability $p(c \mid \mathbf{x})$ be estimated without imposing any model assumptions, using the LSPC formulation [49, 50]. Using this formulation $p(c \mid \mathbf{x})$ of each class c having N_c training samples can be modeled as a linear combination of a large

number of kernel functions as [49, 50]

$$r(c \mid \mathbf{x}; \boldsymbol{\gamma}^c) = \sum_{l=1}^{N_c} \gamma_l^c K(\mathbf{x}, \mathbf{x}_l^c) = (\boldsymbol{\gamma}^c)^{\mathrm{T}} \mathbf{K}^c(\mathbf{x}) \tag{5.6}$$

where $r(c \mid \mathbf{x}; \boldsymbol{\gamma}^c)$ is a multi-kernel approximation of $p(c \mid \mathbf{x})$, T represents the transpose of a vector or matrix, $\mathbf{x}_l^c$ is the lth training sample in training set of class c, $\boldsymbol{\gamma}^c = \left[\gamma_1^c, \gamma_2^c, \ldots, \gamma_{N_c}^c\right]^{\mathrm{T}}$is a vector containing the parameters to be learned from training set, $\mathbf{K}^c(\mathbf{x})$ which is a vector of kernel functions $K(\mathbf{x}, \mathbf{x}_l^c)$ (Gaussian kernel with width β). The Gaussian kernel can be written as [49, 50]

$$K(\mathbf{x}, \mathbf{x}') = \exp(-\beta \|\mathbf{x} - \mathbf{x}'\|^2). \tag{5.7}$$

The parameters vector $\boldsymbol{\gamma}^c$ can be estimated by minimizing a cost function based on expectation of square of error between the modeled and actual class posterior probability of $\mathbf{x}$ over the unknown density $p(\mathbf{x})$ of the target class c. The cost function can be written as [49, 50]

$$J(\boldsymbol{\gamma}^c) = 1/2 \int_{\mathbf{x}_1^c}^{\mathbf{x}_{N_c}^c} \left[r(c \mid \mathbf{x}; \boldsymbol{\gamma}^c) - p(c \mid \mathbf{x})\right]^2 p(\mathbf{x}) d\mathbf{x}. \tag{5.8}$$

By definition $p(c \mid \mathbf{x}) = p(\mathbf{x},c)/p(\mathbf{x})$ which can be replaced in (5.8) as

$$J(\boldsymbol{\gamma}^c) = 1/2 \int_{\mathbf{x}_1^c}^{\mathbf{x}_{N_c}^c} r(c \mid \mathbf{x}; \boldsymbol{\gamma}^c)^2 p(\mathbf{x}) d\mathbf{x} - \int_{\mathbf{x}_1^c}^{\mathbf{x}_{N_c}^c} r(c \mid \mathbf{x}; \boldsymbol{\gamma}^c) p(\mathbf{x}, c) d\mathbf{x} + 1/2 \int_{\mathbf{x}_1^c}^{\mathbf{x}_{N_c}^c} p(\mathbf{x},c)^2/\mathbf{p}(\mathbf{x}) \mathbf{d}\mathbf{x}. \tag{5.9}$$

The third term in (5.9) is a constant. By substituting the values from (5.6), (5.9) gets the

form

$$J(\gamma^c) = 1/2 \int_{\mathbf{x}_1^c}^{\mathbf{x}_{N_c}^c} \left[(\gamma^c)^{\mathrm{T}} \mathbf{K}^c(\mathbf{x}) \right]^2 p(\mathbf{x}) d\mathbf{x} - \int_{\mathbf{x}_1^c}^{\mathbf{x}_{N_c}^c} \left[(\gamma^c)^{\mathrm{T}} \mathbf{K}^c(\mathbf{x}) \right] \mathbf{p}(\mathbf{x}, \mathbf{c}) \mathbf{d}\mathbf{x} + \text{constant}, \quad (5.10)$$

which can be rearranged into a system of linear equations in matrix form as [49, 50]

$$\begin{aligned} J(\gamma^c) = \; & 1/2 (\gamma^c)^{\mathrm{T}} \left[\int_{\mathbf{x}_1^c}^{\mathbf{x}_{N_c}^c} \mathbf{K}^c(\mathbf{x}) \mathbf{K}^c(\mathbf{x})^{\mathrm{T}} p(\mathbf{x}) d\mathbf{x} \right] \gamma^c - \left[\int_{\mathbf{x}_1^c}^{\mathbf{x}_{N_c}^c} \mathbf{K}^c(\mathbf{x}) p(\mathbf{x}, c) d\mathbf{x} \right]^{\mathrm{T}} \gamma^c + \text{constant} \\ & = 1/2 (\gamma^c)^{\mathrm{T}} \mathbf{H}^c \gamma^c + (\mathbf{h}^c)^{\mathrm{T}} \gamma^c + \text{constant}, \end{aligned} \quad (5.11)$$

where $\mathbf{H}^c$ and $\mathbf{h}^c$ are $N_c \times N_c$ matrix and N_c dimensional vector, respectively, given by [49, 50]

$$\begin{aligned} \mathbf{H}^c &= \int_{\mathbf{x}_1^c}^{\mathbf{x}_{N_c}^c} \mathbf{K}^c(\mathbf{x}) \mathbf{K}^c(\mathbf{x})^{\mathrm{T}} p(\mathbf{x}) d\mathbf{x} \\ \mathbf{h}^c &= \int_{\mathbf{x}_1^c}^{\mathbf{x}_{N_c}^c} \mathbf{K}^c(\mathbf{x}) p(\mathbf{x}, c) d\mathbf{x}. \end{aligned} \quad (5.12)$$

After introducing sample averages over the target class c as approximations of the expectations over unknown densities $p(\mathbf{x})$ and $p(\mathbf{x}, c)$, (5.12) can be rewritten as [49, 50]

$$\begin{aligned} \widehat{\mathbf{H}}^c &= 1/N_c \sum_{j=1}^{N_c} \mathbf{K}^c(\mathbf{x}_j^c) \mathbf{K}^c(\mathbf{x}_j^c)^{\mathrm{T}} \\ \widehat{\mathbf{h}}^c &= 1/N_c \sum_{j=1}^{N_c} \mathbf{K}^c(\mathbf{x}_j^c). \end{aligned} \quad (5.13)$$

The vector of unknown parameters γ^c can be found by solving the following regularized optimization problem with $\alpha > 0$ as regularization parameter which is selected appropriately by 10-fold cross validation [49, 50]

$$\gamma^c = \underset{\gamma^c \in \mathbb{R}^{N_c}}{\operatorname{argmin}} \left[1/2 \left(\gamma^c\right)^{\mathrm{T}} \widehat{\mathbf{H}}^c \gamma^c + \left(\widehat{\mathbf{h}}^c\right)^{\mathrm{T}} \gamma^c + \alpha \left(\gamma^c\right)^{\mathrm{T}} \gamma^c \right]. \tag{5.14}$$

By taking derivative of the objective function in (5.14) with respect to γ^c and equating to 0, a system of linear equations is formed which can be solved for γ^c as [49, 50]

$$\begin{aligned} &\left(\widehat{\mathbf{H}}^c + \alpha \mathbf{I}_{N_c}\right) \gamma^c = \widehat{\mathbf{h}}^c \\ &\implies \gamma^c = \left(\widehat{\mathbf{H}}^c + \alpha \mathbf{I}_{N_c}\right)^{-1} \widehat{\mathbf{h}}^c. \end{aligned} \tag{5.15}$$

In order to ensure that the estimated class posterior probabilities are positive, the negative values in vector γ^c are replaced with 0s as [49, 50]

$$\widehat{\gamma}^c = \max\left(0_{N_c}, \gamma^c\right), \tag{5.16}$$

where 0_{N_c} is an N_c dimensional vector and the "max" operation is performed element-wise between the two vectors. The class posterior probability of a test sample $\widetilde{\mathbf{x}}$ for class c can then be estimated as [49, 50]

$$q(c|\mathbf{x} = \widetilde{\mathbf{x}}) = \sum_{l=1}^{N_c} \widehat{\gamma}_l^c \exp\left(-\beta \|\widetilde{\mathbf{x}} - \mathbf{x}_l^c\|^2\right), \tag{5.17}$$

where parameter β is selected by cross validation. To ensure that the estimated class posterior probability is in fact a conditional probability, it is normalized as [49, 50]

$$\widehat{p}(c|\mathbf{x} = \widetilde{\mathbf{x}}) = \frac{\left[\sum_{l=1}^{N_c} \widehat{\gamma}_l^c \exp\left(-\beta \|\widetilde{\mathbf{x}} - \mathbf{x}_l^c\|^2\right)\right]}{\left[\sum_{c'=1}^{C} \sum_{l=1}^{N_{c'}} \widehat{\gamma}_l^{c'} \exp\left(-\beta \|\widetilde{\mathbf{x}} - \mathbf{x}_l^{c'}\|^2\right)\right]}. \tag{5.18}$$

For 2-class classification problem i.e. when $C = 2$, the class label can be assigned to a test

sample as

$$\Theta\left(\widetilde{\mathbf{x}}\right) = \begin{cases} 1 & \text{if } \widehat{p}(c=1|\mathbf{x}=\widetilde{\mathbf{x}})/\widehat{p}(c=2|\mathbf{x}=\widetilde{\mathbf{x}}) > 1 \\ 2 & \text{otherwise.} \end{cases} \tag{5.19}$$

Some samples may experience small changes in the class posterior probabilities, but not enough to exceed the class posterior probability of the other class. Therefore, these samples will not be labeled as "change" using the criterion given in (5.19). To this end, thanks to the probabilistic nature of LSPC which can be utilized in tuning a more suitable threshold λ. It can be tuned, along with β and α, using the model selection procedure explained in the next subsection. Using this threshold, the labeling criterion in (5.19) becomes

$$\Theta\left(\widetilde{\mathbf{x}}\right) = \begin{cases} 1 & \text{if } \widehat{p}(c=1|\mathbf{x}=\widetilde{\mathbf{x}})/\widehat{p}(c=2|\mathbf{x}=\widetilde{\mathbf{x}}) > \lambda \\ 2 & \text{otherwise.} \end{cases} \tag{5.20}$$

The training and testing phases of the change detection technique explained above, are summarized in Algorithm 5.1.

5.2.3 Model Selection Procedure

Model selection using 10-fold cross validation is a common way to tune some unknown parameter(s) of a classifier(s) e.g. β, α and λ here. The studies in [49, 50] selected β and α empirically from pre-specified lists of values, using 10-fold cross validation. This way of model selection is often based on trail and error selection of the list of parameter values. The model parameters selected in this way may be sub-optimal due to the fact that optimal values may not be present in the pre-specified lists of values. Near-optimal combination of these parameters can be found by minimizing a cost function, based on performance indices calculated for 10-fold cross validation experiment on the training set,

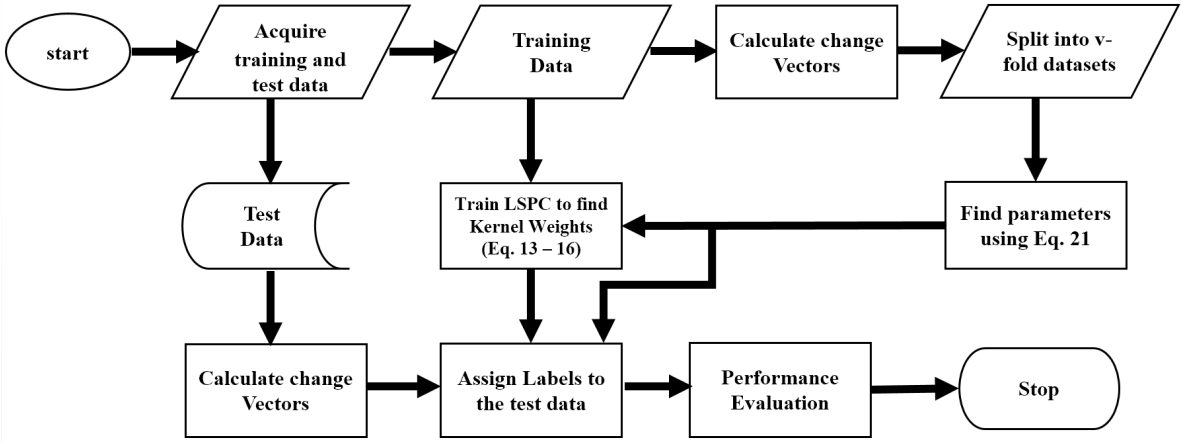

Figure 5.2: Flow chart presenting an overview of different steps involved in the proposed framework.

using an optimization package. The cost function to be minimized can be formulated as

$$\begin{aligned}(\beta,\alpha,\lambda)' &= \underset{\beta,\alpha,\lambda}{\operatorname{argmin}} \left\{ \sqrt{FN_{val}^2 + FP_{val}^2} \right\} \\ &\text{s.t. } FN_{val} < v \geq 0,\ FP_{val} < u \geq 0\end{aligned} \tag{5.21}$$

where FN_{val} and FP_{val} are false positives (false-alarms) and false negatives (miss-alarms, see Table 5.1), respectively, v and u are their respective upper bounds which can be set by the user, and obtained after 10-fold cross validation, $(\beta,\alpha,\lambda)'$ is the vector of the optimized parameter values. The cost function (5.21) can be optimized using any constrained optimization algorithm. In this implementation, MATLAB built in package of Genetic Algorithm (GA) [64] was used for this propose. A number of experiments were carried out for model selection, in order to assess the reliability of the procedure and effect of different parameters of GA. It was observed that these parameters can be tuned conveniently using the proposed model selection procedure. Moreover, changing parameters above the default values set by the MATLAB, did not affect the overall accuracy and performance of the proposed framework. The default MATLAB parameters performed well, hence were left unchanged. Algorithm 5.2 summarizes the steps for evaluating the fitness func-

Algorithm 5.1

Training

1. Given inputs: training set $\mathbf{X}_{tr}$, parameter β, and parameter α
2. For $c = 1, 2, \ldots, C$ do
3. Calculate $\widehat{\mathbf{H}}^c$ using Equation (5.13).
4. Calculate $\widehat{\mathbf{h}}^c$ using Equation (5.13).
5. Determine $\widehat{\gamma}^c$ using Equations (5.15) and (5.16).
6. endfor
7. Output: $\widehat{\gamma}^c$

Testing

1. Given inputs: test sample $\widetilde{\mathbf{x}}$, parameters $\widehat{\gamma}^c$, β, α and threshold λ.
2. Use Equation (5.18) to determine its $\widehat{p}(c|\mathbf{x} = \widetilde{\mathbf{x}})$.
3. Use Equation (5.20) to assign it a class label.
4. Output: class label $\Theta(\widetilde{\mathbf{x}})$

tion of GA expressed in (5.21). An overview of different steps involved in the proposed framework are shown in the flowchart in Figure 5.2.

5.3 Results and Discussion

The performance of the proposed method was evaluated and compared to widely used Naive Bayes (NB) classifier with Gaussian assumption and Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel for detection of defoliation, in *Eucalyptus nitens* and *E. globulus plantations* caused by leaf beetles (*Paropsisterna* spp), using Landsat 7 ETM+ multi-date images. All the methods were tested on raw bands and band indices data sets of both the areas in order to check their robustness for different data sets with difference between their data distributions. For implementation of SVM, LIBSVM Mat-

Algorithm 5.2 $\mathbf{X}_{tr}$, β and α

1. Set $\mathbf{Z}_{val} \longleftarrow \mathbf{X}_{tr}$.
2. Set $FN_{val}, FP_{val} = 0$.
3. Divide $\mathbf{Z}_{val}$ into 10 equal sized subsets: $\mathbf{Z}_{val} = \left\{\mathbf{Z}_{val}^{g}\right\}_{g=1}^{10}$.
4. For $i = 1$ to 10
5. Set training set $\mathbf{X}'_{tr}$= $\left\{\mathbf{Z}_{val}^{g}\right\}_{g=1, g\neq i}^{10}$.
6. Set test set $\mathbf{X}'_{te} = \mathbf{Z}_{val}^{i} = \left\{\widetilde{\mathbf{x}}_{te}^{k}\right\}_{k=1}^{N_{tr}/10}$.
7. Use Algorithm (5.1) to find class labels set: $\Theta_{val}^{i} = \left\{\Theta\left(\widetilde{\mathbf{x}}_{te}^{k}\right)\right\}_{k=1}^{N_{tr}/10}$
8. endfor
9. Calculate FN_{val} and FP_{val} from $\Theta_{val} = \left\{\Theta_{val}^{k}\right\}_{k=1}^{10}$.
10. Evaluate the cost function using Equation (5.21).
11. Output: $\mathscr{L}(\beta, \alpha, \lambda)$

lab library was used, which is freely available on the World Wide Web (WWW) [65]. For model selection, one of the most popular methods used in the literature is grid search method [66–68]. The LIBSVM has a built-in option of cross validation, which was used (10-fold) in grid search method to select optimal parameter values for the SVM used in this study.

Figure 5.3 shows the change detection results of the proposed method on raw bands and band indices data sets of both the study areas, along with their respective reference maps (ground truth maps, containing 8764 and 8326 infested/defoliated pixels in area 1 and 2, respectively). The top row of the figure presents the reference maps, the second row presents the change detection map produced on the raw bands data sets and the third row shows the change maps produced on the band indices data sets. The first and second columns of the figure represent Area 1 and Area 2, respectively. Figure 5.4 shows the miss-classifications (errors i.e. FN and FP) incurred by the proposed method. All the

Table 5.1: Definitions of performance indices used in this study. $\Theta(\widetilde{\mathbf{x}})$=classifier labeling function, $g(\widetilde{\mathbf{x}})$ =ground truth label, n_c =number of change samples, n_{nc} =number of no-change samples.

Performance Index	Definition	Relationship
TP (%)	$p(\Theta(\widetilde{\mathbf{x}}) = 1 \mid g(\widetilde{\mathbf{x}}) = 1) \times 100$	TP+FN=100
FN (%)	$p(\Theta(\widetilde{\mathbf{x}}) = 2 \mid g(\widetilde{\mathbf{x}}) = 1) \times 100$	
TN (%)	$p(\Theta(\widetilde{\mathbf{x}}) = 2 \mid g(\widetilde{\mathbf{x}}) = 2) \times 100$	TN+FP=100
FP (%)	$p(\Theta(\widetilde{\mathbf{x}}) = 1 \mid g(\widetilde{\mathbf{x}}) = 2) \times 100$	
OA (%)	$((\mathrm{TP} \times n_c) + (\mathrm{TN} \times n_{nc})) / (n_c + n_{nc})$	-

maps in Figure 5.4 correspond to their respective change maps in Figure 5.3. The miss-classification images contain those points which were originally labeled as changed in ground truth data but detected as no-change by the algorithm or vice versa. It can be observed from Figures 5.3 and 5.4 that the proposed method detects changes in both raw bands and band indices data sets of both the study areas accurately and robustly. The larger connected areas in the miss-classification maps corresponding to Area 2, are FPs, which can be attributed to labeling inaccuracies incurred by the surveyors.

The figures have been presented only for visual illustration and may not be very informative about the change detection performance of the proposed method. Therefore, its quantitative results and their comparison with two well known classifiers, namely NB and SVM, have been summarized in Tables 5.2, 5.3 and 5.4 for further assessment. Table 5.2 and 5.3 present the performance comparison in terms performance indices True Positives (TP, correct alarms), True Negatives (TN, correct no-alarms), Overall Accuracy (OA) and kappa-statistic (κ) (see Table 5.1 for the definitions of the performance indices used), for Area 1 and 2, respectively. Despite some criticism on using kappa-statistic for accuracy assessment, it is still one of the most popular performance metrics used [69]. Table 5.4 presents the corresponding McNemar's test [70] which quantifies significance of the differences of the proposed method with NB and SVM. The tables show that the LSPC based method performs robustly in all the cases. It clearly outperforms the NB classifier on all

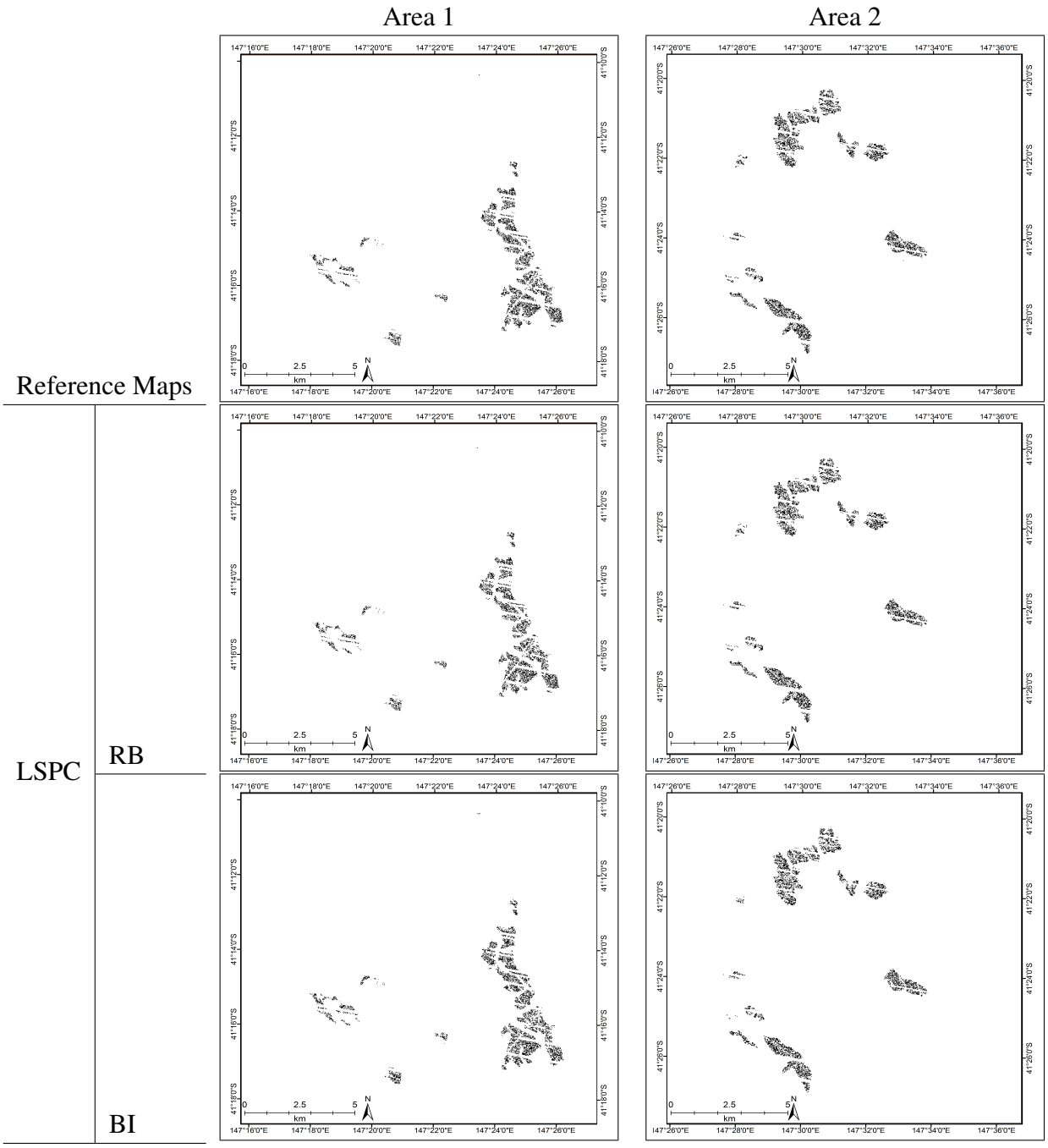

Figure 5.3: Changes detected (TP) by the proposed method in Raw Bands (RB) and Band Indices (BI) data sets of Area 1 and Area 2, along with reference maps (top row).

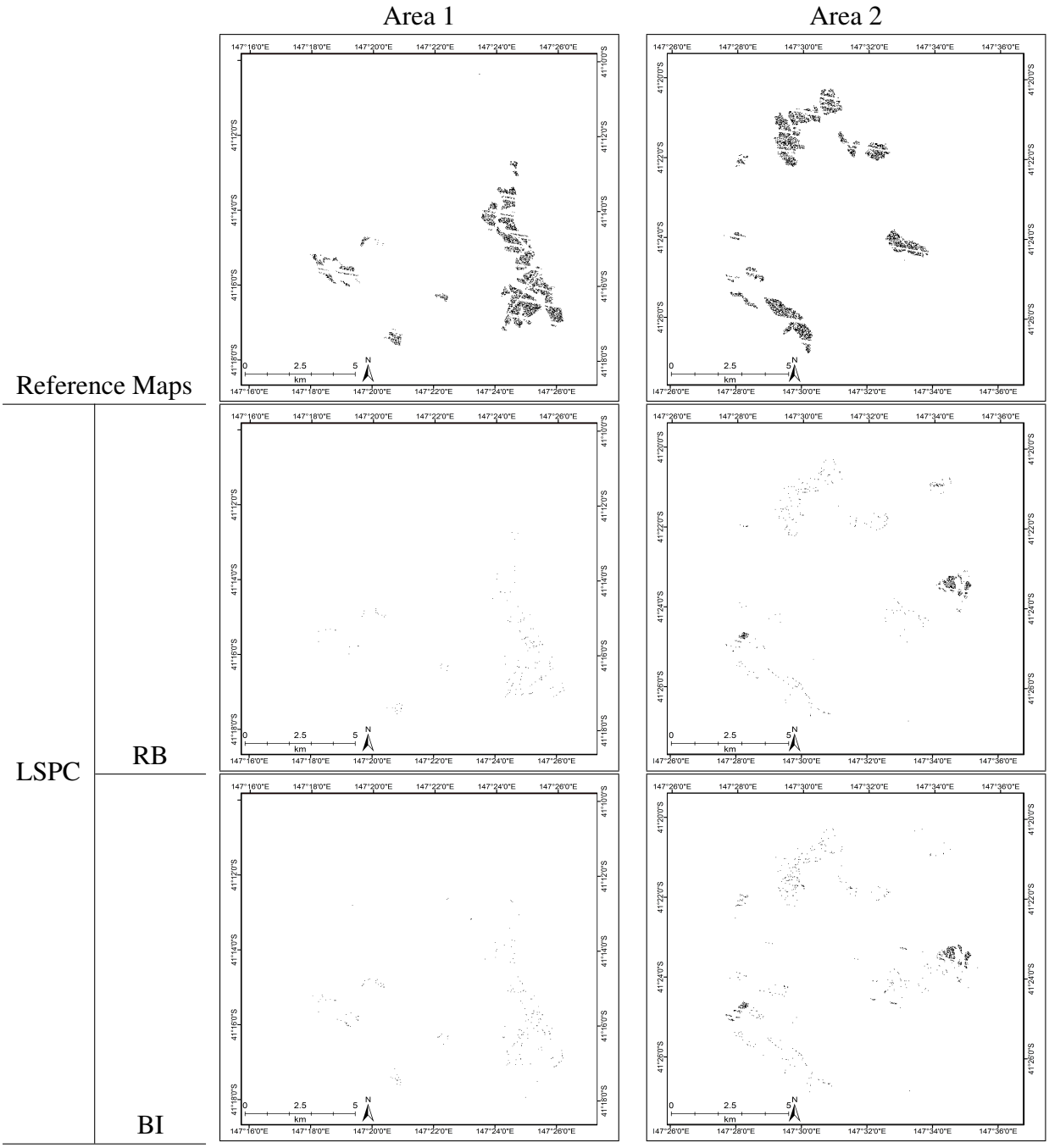

Figure 5.4: Miss-classifications (FN and FP) by the proposed method in Raw Bands (RB) and Band Indices (BI) data sets of Area 1 and Area 2, along with reference maps (top row).

the data sets, with OA ranging from 93.6% ($\kappa = 0.87$) for Area 2 to 97.4% ($\kappa = 0.94$) for Area 1, against 85.3% ($\kappa = 0.69$) to 93.4% ($\kappa = 0.85$), respectively. The corresponding McNemar's tests in Table 5.4 also show that these differences in OA of both the methods are statistically significant for all the data sets. This difference in the performances can be observed in Figure 5.5 which gives a comparison between the Receiver Operating Characteristics (ROC) curves of the two methods, drawn between different pairs of TP and FP achieved by varying their thresholds, for both types of data sets of both the test areas considered here. Considering the comparison between LSPC and SVM, the SVM's OA ranges from 93.4% ($\kappa = 0.86$) in case of Area 2, to 97.4% ($\kappa = 0.94$) in case of Area 1, which is almost the same as that of the LSPC as confirmed by the corresponding McNemar's tests in Table 5.4.

It can be noticed from the results that both the approaches with non-parametric models, that are free of model assumptions, are more robust and accurate on different data sets with unknown distributions as compared to the Gaussian based NB classifier. This can be attributed to the fact that the distributions of the data sets are unknown, and are very rarely Gaussian in practical cases, hence violates the assumptions made in NB and resulting more miss-classifications. Please note that the McNemar's test only tells whether the absolute number of samples, on which there is a disagreement between the two classifiers, is significant or not. It neither tells which classifier has the better performance nor does it take the data set size into consideration. Therefore, the OA accuracy comparisons and the McNemar's test results should be considered simultaneously in order to draw meaningful conclusions. In Table 5.4 the McNemar's test was conducted after scaling the data set size and the observed results to out-of-1000 equivalents.

These findings are consistent with many other studies in which non-parametric kernel based methods outperformed the parametric approaches, e.g. a recently published study [6] demonstrated improvement in performance of the change detection algorithm by using

Repeated Sequential Probability Ratio Test (RSPRT) based on relative density ratios [51, 52], estimated directly from the data by non-parametric multi-kernel approach [51–54], over the traditional likelihood ratios based on Gaussian assumption. Studies in [71–74] showed superior classification performance of non-parametric kernel based methods over parametric Bayesian classifiers.

The model parameters i.e. kernel width β, regularization parameter α and threshold λ, were tuned by the model selection algorithm proposed here. Since fractional changes ($< 1\%$) in the false negative and false positive values do not have any significant importance in this case, the model selection procedure can be stopped after a suitable number of iterations, without waiting for fractional improvements in the cost function. Model parameters values corresponding to the iteration with the lowest value of the cost function are chosen. In this study, 50 iterations proved to be sufficient for model selection. The comparison between the accuracies and the corresponding (α, β) parameters, achieved by LSPC using simple cross validation (L2) and LSPC with the proposed model selection algorithm (L1), is given in Table 5.5. The significance of the results presented in Table 5.5 has been calculated through McNemar's tests shown in Table 5.6. It can be observed that there is a significant improvement in the accuracy when the proposed method is used with the proposed model selection procedure, over when the proposed method is used with parameters selected by 10-fold cross validation. The improvement in the accuracy ranged from 2.6% to 4.9%.

5.4 Conclusions

In this study, a non-parametric algorithm was proposed for detecting defoliation caused by leaf beetle in *Eucalyptus Nitens* and *E. Globulus* plantations. To this end, a new non-parametric approach LSPC was applied for finding optimal decision boundary between change and no-change Change Vectors (CVs) calculated after image differencing. An

Table 5.2: Results Comparison between LSPC (Least Squares Probabilistic Classifier), NB (Naive Bayes) Classifier and SVM (Support Vector Machine), in terms of performance metrics, for Area 1. TP= True Positives, TN= True Negatives, OA = Overall Accuracy, κ= Kappa-Statistic.

Area 1												
Features	**Raw Bands**						**Band Combination Indices**					
Algorithm	**TP**	**FN**	**TN**	**FP**	**OA**	κ	**TP**	**FN**	**TN**	**FP**	**OA**	κ
LSPC	**99.8**	**0.2**	**92.0**	**8.0**	**97.4**	**0.94**	**99.3**	0.7	**91.6**	8.4	**96.9**	**0.93**
NB	94.6	5.4	92.1	7.9	93.4	0.85	90.8	9.2	90.6	9.4	90.7	0.79
SVM	99.8	0.2	91.9	8.1	97.4	0.94	99.6	0.4	91.4	8.6	97.1	0.93

Table 5.3: Results Comparison between LSPC (Least Squares Probabilistic Classifier), NB (Naive Bayes) Classifier and SVM (Support Vector Machine), in terms of performance metrics, for Area 2. TP= True Positives, TN= True Negatives, OA = Overall Accuracy, κ= Kappa-Statistic.

Area2												
Features	**Raw Bands**						**Band Combination Indices**					
Algorithm	**TP**	**FN**	**TN**	**FP**	**OA**	κ	**TP**	**FN**	**TN**	**FP**	**OA**	κ
LSPC	**98.2**	1.8	**90.8**	9.2	**93.6**	**0.87**	**97.6**	2.4	**91.5**	8.5	**93.7**	**0.87**
NB	90.4	9.6	82.4	17.6	85.3	0.69	88.3	11.7	84.2	15.8	85.7	0.70
SVM	97.3	2.7	91.3	9.7	93.5	0.86	96.3	3.7	91.8	8.2	93.4	0.86

automated model selection approach was also designed for near-real optimal selection of the model parameters. It was found that the Landsat imagery suits the detection of defoliation caused by leaf beetles (*Paropsisterna* spp) in *Eucalyptus Nitens* and *E. Globulus* plantations. Furthermore, the LSPC can model the class posterior probabilities robustly. Its comparison with the parametric NB and SVM suggests that it performs comparably to SVM in all the cases and is more robust to variations in data as compared to the NB. Furthermore, it can be concluded that the proposed model selection strategy has an advantage over the traditional model selection procedure which is based on trial and error selection of the list of parameter values to be tuned, hence its accuracy being dependent on goodness of the list of parameter values.

While Bayesian Classifiers have the advantage of multi-class capabilities and probabilis-

Table 5.4: Results Comparison between LSPC (Least Squares Probabilistic Classifier), NB (Naive Bayes) Classifier and SVM (Support Vector Machine), in terms of McNemar's test statistics (χ^2-values with 1 degree of freedom and 95% confidence). $p < 0.05$ columns indicate whether the test statistic has probability of less than 0.05 (outside 95% confidence interval) or not, Signf. = Significance.

Method	Area 1						Area 2					
	Raw Bands			Band Indices			Raw Bands			Band Indices		
	χ^2	$p < 0.05$	Signif.	χ^2	$p < 0.05$	Signif.	χ^2	$p < 0.05$	Signif.	χ^2	$p < 0.05$	Signif.
NB	24.9	yes	yes	52.6	yes	yes	74.5	yes	yes	48.9	yes	yes
SVM	0.5	no	no	1.1	no	no	0.001	no	no	2.3	no	no

Table 5.5: Comparison between accuracies and the corresponding (α, β) parameters, achieved by LSPC with the proposed model selection (denoted as L1) and with the simple 10-fold cross validation (denoted as L2). Algo. = Algorithm, OA = Overall Accuracy, κ= Kappa-Statistic.

Area	Area 1				Area 2			
	Raw Bands		Band Indices		Raw Bands		Band Indices	
Algo.	**OA**	κ	**OA**	κ	**OA**	κ	**OA**	κ
L1	**97.4**	**0.94**	**96.9**	**0.93**	**93.6**	**0.87**	**93.7**	**0.87**
L2	94.8	0.88	93.8	0.86	90.7	0.81	88.8	0.77

Table 5.6: Results Comparison between LSPC with the proposed model selection (denoted as L1) and with the simple 10-fold cross validation (denoted as L2), in terms of McNemar's test statistics (χ^2-values with 1 degree of freedom and 95% confidence). $p < 0.05$ columns indicate whether the test statistic has probability of less than 0.05 (outside 95% confidence interval) or not, Signf. = Significance.

Area 1						Area 2					
Raw Bands			Band Indices			Raw Bands			Band Indices		
χ^2	$p < 0.05$	Signif.	χ^2	$p < 0.05$	Signif.	χ^2	$p < 0.05$	Signif.	χ^2	$p < 0.05$	Signif.
27.3	yes	yes	43.8	yes	yes	38.1	yes	yes	50.7	yes	yes

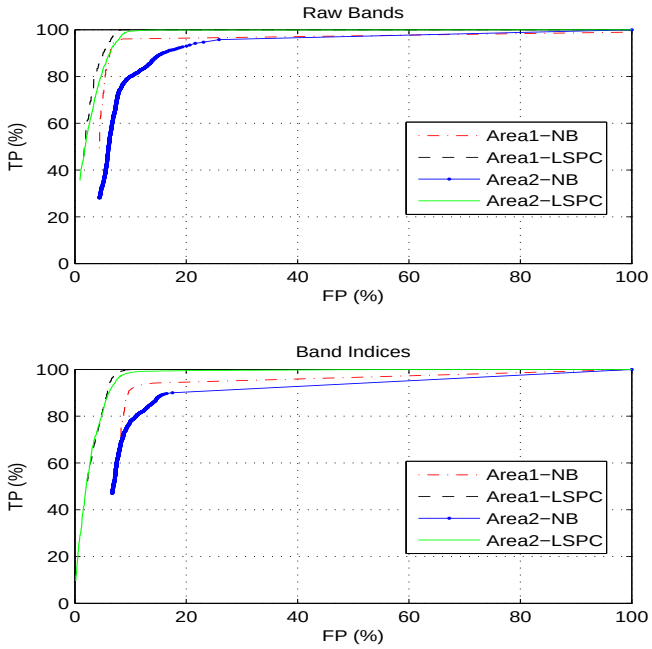

Figure 5.5: Comparison between ROC curves of LSPC and NB (Naive Bayes). Top: ROC comparison on raw bands data sets of both the areas. Bottom: ROC comparison on band indices data sets of both the areas.

tic nature, their bottleneck is their dependence on the model assumption which may be far from the original data distribution and hence can limit their performance, as observed in this study. Similarly, SVM has the advantage of being non-parametric, however it is basically a binary classifier which assigns only class labels (sign) to a test sample, hence needing different schemes to be designed for it in order to handle multi-class problems and provide probabilistic output. The facts that LSPC is not only non-parametric as SVM, but also probabilistic in nature, capable of handling multi-class problems, easy to implement (thanks to its closed form analytically tractable solution) and has detection accuracy comparable to SVM, make it interesting and promising for the remote sensing community, especially when class memberships of the samples in multi-class problems need to be identified. Although, the computational efficiency was not crucial, hence not considered here, nevertheless, it is worth mentioning here for potential readers, for whom computational efficiency may be crucial, that the LSPC was shown to be computationally more

efficient than SVM for multi-class problems [49, 50]. However, its superiority over SVM, in terms of computational efficiency, is yet to be proven in remote sensing applications.

The LSPC proved useful when used as a part of a change detection framework in a 2-class classification problem. However, its multi-class performance on remote sensing data sets is still an open question for future research. Being of statistical nature and independent of the type of data, it can be applied to a wide range of classification applications of EO data as its scope is not limited to vegetation data only. Moreover, apart from classification applications, it can also be applied for non-parametric approximation of conditional distribution of the data in many other applications.

Acknowledgments

The authors are thankful to the Forestry Tasmania for providing the shape files of the ground truth data, Faheem Iqbal and Irfan Akhter Iqbal (School of Land and Food, GIS group, University of Tasmania) for providing help in using ArcGIS, and B.P. Salmon (School of Engineering, University of Tasmania, Australia) for his helpful comments during the development of this manuscript.

Bibliography

[1] A. Singh, "Review article digital change detection techniques using remotely-sensed data," *International Journal of Remote Sensing*, vol. 10, no. 6, pp. 989–1003, 1989.

[2] P. Coppin, I. Jonckheere, K. Nackaerts, B. Muys, and E. Lambin, "Digital change detection methods in ecosystem monitoring: a review," *International Journal of Remote Sensing*, vol. 25, no. 9, pp. 1565–1596, 2004.

[3] D. Lu, P. Mausel, E. Brondízio, and E. Moran, "Change detection techniques," *International Journal of Remote Sensing*, vol. 25, no. 12, pp. 2365–2401, 2004.

[4] A. Anees and J. Aryal, "Near-real time detection of beetle infestation in pine forests using modis data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, pp. 3713–3723, Sept 2014.

[5] A. Anees and J. Aryal, "A statistical framework for near-real time detection of beetle infestation in pine forests using modis data," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, pp. 1717–1721, Oct 2014.

[6] A. Anees, J. Aryal, M. O'Reilly, and T. Gale, "A relative density ratio-based framework for detection of land cover changes in modis ndvi time series," *IEEE Journal ofSelected Topics in Applied Earth Observations and Remote Sensing*, vol. PP, no. 99, pp. 1–13, 2015.

[7] M. A. Wulder, C. C. Dymond, J. C. White, D. G. Leckie, and A. L. Carroll, "Sur-

veying mountain pine beetle damage of forests: A review of remote sensing opportunities," *Forest Ecology and Management*, vol. 221, no. 1, pp. 27–41, 2006.

[8] M. Wulder, J. White, B. Bentz, M. Alvarez, and N. Coops, "Estimating the probability of mountain pine beetle red-attack damage," *Remote Sensing of Environment*, vol. 101, pp. 150–166, Mar 2006.

[9] A. Meddens, J. Hicke, L. Vierling, and A. Hudak, "Evaluating methods to detect bark beetle-caused tree mortality using single-date and multi-date landsat imagery," *Remote Sensing of Environment*, vol. 132, pp. 49–58, May 2013.

[10] G. Amman and W. Cole, "Mountain pine beetle dynamics in lodgepole pine forests. part ii: population dynamics.," *Gen. Tech. Rep. INT-146. Ogden, UT:U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station*, 1983.

[11] S. Embrey, J. V. Remais, and H. J., "Climate change and ecosystem disruption: the health impacts of the north american rocky mountain pine beetle infestation," *American Journal of Public Health*, vol. 5, pp. 818–827, May 2012.

[12] Y. Mukai, T. Sugimura, H. Watanebe, and K. Wakamori, "Extraction of areas infested by pine bark beetle using landsat mss data," *Photogrammetric Engineering and Remote Sensing*, vol. 53, no. 1, pp. 77–81, 1987.

[13] D. Werle, Y. Lee, and R. Brown, "The use of integrated multispectral and radar remote sensing data for monitoring forest clearcut and regeneration sites on vancouver island," in *Proceedings of the Tenth Canadian Symposium on Remote Sensing*, (Edmonton, Ottawa, Canada), pp. 319–332, May 5-8 1986.

[14] A. Banner and T. Lynham, "Multitemporal analysis of landsat data for forest cutover mapping: a trial of two procedures," *Proceedings of the 7th Canadian symposium on Remote Sensing*, pp. 233–240, September 8-11 1981.

[15] T. Fung, "An assessment of tm imagery for land cover change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, pp. 681–684, July 1990.

[16] P. Howrath and G. M. Wickware, "Procedures for change detection using landsat digital data," *International Journal of Remote Sensing*, vol. 2, no. 3, pp. 277–291, 1981.

[17] R. D. Johnson and E. S. Kasischke, "Change vector analysis: A technique for the multispectral monitoring of land cover and condition," *International Journal of Remote Sensing*, vol. 19, no. 3, pp. 411–426, 1998.

[18] J. Vogelmann, "Detection of forest change in the green mountains of vermont using multispectral scanner data," *International Journal of Remote Sensing*, vol. 9, no. 7, pp. 1187–1200, 1988.

[19] T. Fung and E. LeDrew, "Application of principal components analysis to change detection," *Photogrammetric Engineering and Remote Sensing*, vol. 53, no. 12, pp. 1649–1658, 1987.

[20] P. Coppin and M. Bauer, "Processing of multitemporal landsat tm imagery to optimize extraction of forest cover change features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, pp. 918–927, Jul 1994.

[21] L. Bruzzone and D. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, pp. 1171–1182, May 2000.

[22] R. Shah-Hosseini, S. Homayouni, and A. Safari, "A hybrid kernel-based change detection method for remotely sensed data in a similarity space," *Remote Sensing*, vol. 7, no. 10, p. 12829, 2015.

[23] S. Gopal and C. Woodcock, "Remote sensing of forest change using artificial neural

networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, pp. 398–404, Mar 1996.

[24] C. Woodcock, S. Macomber, M. Pax-Lenney, and W. Cohen, "Monitoring large areas for forest change using landsat: Generalization across space, time and landsat sensors," *Remote Sensing of Environment*, vol. 78, pp. 194 – 203, October 2001. Landsat 7.

[25] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Rojo-Alvarez, and M. Martinez-Ramon, "Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, pp. 1822–1835, June 2008.

[26] H. Nemmour and Y. Chibani, "Multiple support vector machines for land cover change detection: An application for mapping urban extensions," *{ISPRS} Journal of Photogrammetry and Remote Sensing*, vol. 61, no. 2, pp. 125 – 133, 2006.

[27] D. C. Zanotta and V. Haertel, "Gradual land cover change detection based on multi-temporal fraction images," *Pattern Recogn.*, vol. 45, pp. 2927–2937, Aug. 2012.

[28] R. A. Redner and H. F. Walker, "Mixture densities, maximum likelihood and the em algorithm," *SIAM review*, vol. 26, no. 2, pp. 195–239, 1984.

[29] P. Lombardo and C. Oliver, "Maximum likelihood approach to the detection of changes between multitemporal sar images," *IEE Proceedings on Radar, Sonar and Navigation,*, vol. 148, pp. 200–210, Aug 2001.

[30] I. Rish, "An empirical study of the naive bayes classifier," tech. rep., 2001.

[31] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," *Mach. Learn.*, vol. 29, pp. 131–163, Nov. 1997.

[32] G. Yan, J. F. Mas, B. H. P. Maathuis, Z. Xiangmin, and P. M. Van Dijk, "Comparison of pixel-based and object-oriented image classification approaches- a case study

in a coal fire area, wuda, inner mongolia, china," *International Journal of Remote Sensing*, vol. 27, no. 18, pp. 4039–4055, 2006.

[33] J. Otukei and T. Blaschke, "Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms," *International Journal of Applied Earth Observation and Geoinformation*, vol. 12, pp. S27 – S31, February 2010.

[34] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," *{ISPRS} Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 3, pp. 247 – 259, 2011.

[35] A. Erener, "Classification method, spectral diversity, band combination and accuracy assessment evaluation for urban feature detection," *International Journal of Applied Earth Observation and Geoinformation*, vol. 21, no. 0, pp. 397 – 408, 2013.

[36] C. Huang, L. Davis, and J. Townshend, "An assessment of support vector machines for land cover classification," *International Journal of Remote Sensing*, vol. 23, no. 4, pp. 725–749, 2002.

[37] B. Dixon and N. Candade, "Multispectral landuse classification using neural networks and support vector machines: one or the other, or both?," *International Journal of Remote Sensing*, vol. 29, no. 4, pp. 1185–1206, 2008.

[38] M. Linderman, J. Liu, J. Qi, L. An, Z. Ouyang, J. Yang, and Y. Tan, "Using artificial neural networks to map the spatial distribution of understory bamboo from remote sensing data," *International Journal of Remote Sensing*, vol. 25, pp. 1685–1700, 2004.

[39] D. L. Civco, "Artificial neural networks for land-cover classification and mapping," *International Journal of Geographical Information Systems*, vol. 7, no. 2, pp. 173–186, 1993.

[40] H. Bischof, W. Schneider, and A. Pinz, "Multispectral classification of landsat-images using neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, pp. 482–490, May 1992.

[41] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag New York, Inc., 1995.

[42] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, pp. 273–297, Sept. 1995.

[43] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discov.*, vol. 2, pp. 121–167, June 1998.

[44] T. Kavzoglu and I. Colkesen, "A kernel functions analysis for support vector machines for land cover classification," *International Journal of Applied Earth Observation and Geoinformation*, vol. 11, no. 5, pp. 352 – 359, 2009.

[45] M. Volpi, D. Tuia, G. Camps-Valls, and M. Kanevski, "Unsupervised change detection with kernels," *Geoscience and Remote Sensing Letters, IEEE*, vol. 9, pp. 1026–1030, Nov 2012.

[46] R. Shah-Hosseini, S. Homayouni, and A. Safari, "Environmental monitoring based on automatic change detection from remotely sensed data: kernel-based approach," *Journal of Applied Remote Sensing*, vol. 9, no. 1, pp. 095–992, 2015.

[47] F. Bovolo, G. Camps-Valls, and L. Bruzzone, "A support vector domain method for change detection in multitemporal images," *Pattern Recognition Letters*, vol. 31, no. 10, pp. 1148 – 1154, 2010. Pattern Recognition in Remote SensingFifth {IAPR} Workshop on Pattern Recognition in Remote Sensing (PRRS 2008).

[48] F. Bovolo, L. Bruzzone, and M. Marconcini, "An unsupervised change detection technique based on bayesian initialization and semisupervised svm," in *Geo-*

science and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International, pp. 2370–2373, July 2007.

[49] M. Sugiyama and J. Simm, "A computationally-efficient alternative to kernel logistic regression," in *IEEE International Workshop on Machine Learning for Signal Processing (MLSP), 2010*, pp. 124–129, Aug 2010.

[50] M. Sugiyama, "Superfast-trainable multi-class probabilistic classifier by least-squares posterior fitting," *IEICE Trans. Inf. & Syst.*, vol. 93, pp. 2690–2701, oct 2010.

[51] M. Yamada, T. Suzuki, T. Kanamori, H. Hachiya, and M. Sugiyama, "Relative density-ratio estimation for robust distribution comparison," *Neural computation*, vol. 25, no. 5, pp. 1324–1370, 2013.

[52] S. Liu, M. Yamada, N. Collier, and M. Sugiyama, "Change-point detection in time-series data by relative density-ratio estimation," *Neural Networks*, vol. 43, no. 0, pp. 72 – 83, 2013.

[53] M. Sugiyama, S. Nakajima, H. Kashima, P. Von Bünau, and M. Kawanabe, "Direct importance estimation with model selection and its application to covariate shift adaptation," in *Advances in neural information processing systems*, pp. 1433–1440, 2008.

[54] Y. Kawahara and M. Sugiyama, "Change-point detection in time-series data by direct density-ratio estimation," in *Proceedings of the SIAM International Conference on Data Mining, SDM 2009, April 30 - May 2, 2009, Sparks, Nevada, USA*, pp. 389–400, SIAM, 2009.

[55] X. Chen, L. Vierling, and D. Deering, "A simple and effective radiometric correction method to improve landscape change detection across sensors and across time," *Remote Sensing of Environment*, vol. 98, no. 1, pp. 63–79, 2005.

[56] C. Parente, "Toa reflectance and ndvi calculation for landsat 7 etm+ images of sicily," *Proceedings in EIIC-The 2nd Electronic International Interdisciplinary Conference*, no. 1, 2013.

[57] E. Vermote, D. Tanre, J. Deuze, M. Herman, and J.-J. Morcette, "Second simulation of the satellite signal in the solar spectrum, 6s: an overview," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, pp. 675–686, May 1997.

[58] X. Xiao, Q. Zhang, B. Braswell, S. Urbanski, S. Boles, S. Wofsy, B. Moore, and D. Ojima, "Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data," *Remote Sensing of Environment*, vol. 91, pp. 256–270, May 2004.

[59] J. Rouse, R. Haas, J. Schell, and D. Deering, "Monitoring vegetation systems in the great plains with erts," *NASA Special Publication*, vol. 351, p. 309, 1974.

[60] C. Tucker, "Red and photographic infrared linear combinations for monitoring vegetation," *Remote Sensing of Environment*, vol. 8, no. 2, pp. 127 – 150, 1979.

[61] H. Xu, "Modification of normalised difference water index (ndwi) to enhance open water features in remotely sensed imagery," *International Journal of Remote Sensing*, vol. 27, no. 14, pp. 3025–3033, 2006.

[62] H. Qing Liu and A. Huete, "A feedback based modification of the ndvi to minimize canopy background and atmospheric noise," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, pp. 457–465, Mar 1995.

[63] R. J. Kauth and G. S. Thomas, "The tasseled cap- a graphic description of the spectral-temporal development of agricultural crops as seen by landsat," in *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*, pp. 41–51, 1976.

[64] W. Banzhaf, F. Francone, R. Keller, and P. Nordin, *Genetic Programming: An Intro-*

duction: on the Automatic Evolution of Computer Programs and Its Applications. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1998.

[65] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011.

[66] G. E. Hinton, *Neural Networks: Tricks of the Trade: Second Edition*, ch. A Practical Guide to Training Restricted Boltzmann Machines, pp. 599–619. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[67] H. Larochelle, D. Erhan, A. Courville, J. Bergstra, and Y. Bengio, "An empirical evaluation of deep architectures on problems with many factors of variation," in *Proceedings of the 24th International Conference on Machine Learning*, ICML '07, (New York, NY, USA), pp. 473–480, ACM, 2007.

[68] Y. LeCun, L. Bottou, G. B. Orr, and K. R. Müller, *Neural Networks: Tricks of the Trade*, ch. Efficient BackProp, pp. 9–50. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998.

[69] A. Stein, J. Aryal, and G. Gort, "Use of the bradley-terry model to quantify association in remotely sensed images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 852–856, April 2005.

[70] Q. McNemar, "Note on the sampling error of the difference between correlated proportions or percentages," *Psychometrika*, vol. 12, pp. 153–157, June 1947.

[71] C. Huang, L. S. Davis, and J. R. G. Townshend, "An assessment of support vector machines for land cover classification," *International Journal of Remote Sensing*, vol. 23, no. 4, pp. 725–749, 2002.

[72] M. Pal and P. M. Mather, "Support vector machines for classification in remote sensing," *International Journal of Remote Sensing*, vol. 26, no. 5, pp. 1007–1011, 2005.

[73] D. Tien Bui, B. Pradhan, O. Lofman, and I. Revhaug, "Landslide susceptibility assessment in vietnam using support vector machines, decision tree, and naive bayes models," *Mathematical Problems in Engineering*, vol. 2012, 2012.

[74] Y. Qian, W. Zhou, J. Yan, and L. Li, W.and Han, "Comparing machine learning classifiers for object-based land cover classification using very high resolution imagery," *Remote Sensing*, vol. 7, no. 1, pp. 153–168, 2014.

6 Conclusions and Future Research

6.1 Conclusions

Reliable land cover monitoring and change detection has been shown to be of significant importance for the global community [1–11]. Its importance was discussed in Chapter 1. This thesis presented several methods for detecting changes in land cover using remote sensing data, with a special emphasis on the changes caused by beetle infestations in pine forests of north America and eucalyptus plantations in north-eastern Tasmania, Australia.

Automated beetle infestation monitoring, using remotely sensed data, has been given considerable importance over past few decades because manual monitoring or aerial surveys of land cover changes, spread over large areas, is very time consuming and involves expensive labor [12–22]. Furthermore, manual monitoring or aerial surveys can also be prone to errors depending upon the experience and skills of the experts carrying on the surveys. The existing methods have certain limitations which need to be addressed in order to enhance their performance. Firstly, majority of the existing methods do not detect beetle infestations in a near-real time manner, which is important for the authorities in order to take timely actions to stop the damage. Therefore, near-real time methods with improved accuracy and detection delays are desirable. Secondly, majority of the methods

depend upon threshold selection from the labelled examples [23, 24], which can be hard to collect, hence methods with automatic threshold selection are needed. Thirdly, most of the test statistics used are based on traditional likelihood ratios based on Gaussian distribution assumption [23–25], which may lead to sub-optimal results in case the data is far from Gaussian. Furthermore, the threshold selection approaches existing in the literature tune the threshold by optimizing only the accuracy (TP and TN), which does not suit near-real time methods. Fourth, many bi-temporal change detection methods are based on the popular Bayesian classifiers e.g. Maximum Likelihood Classifier (MLC) and Naive Bayes (NB) class [26–30], which assumes a parametric model (mostly Gaussian) for the data and can have robustness issues with different types of data.

The methods proposed in this thesis tackled the above limitations of the change detection approaches in remote sensing. The satellite data used in this thesis were MODIS and Landsat 7 ETM+ images because of their free availability and suitability for the problems tackled in this thesis. MODIS (500m 8-days) data is hyper-temporal and coarse spatial resolution which suits detection of changes (infestations) spread on wide areas in near-real time, hence was used for near-real time detection of beetle infestations in pine forests of north America. The leaf beetle defoliation in eucalyptus plantations of north-eastern Tasmania, is not very widely spread, hence the moderate resolution Landsat ETM+ 7 data (30m) data was used for its detection. Four different statistical change detection methods were proposed in this thesis, which were meant for certain improvements over the existing methods.

The NLS and FIR filter based methods were discussed in Chapter 2. The NLS based method uses a μ-parameter time series, extracted by fitting a triply modulated cosine function to the hyper-temporal vegetation index time series of MODIS data, to derive log-likelihood ratio based test statistics. The FIR-based method estimates the μ-parameter time series by convolving a low-pass filter with the hyper-temporal MODIS vegetation

index time series and hence is computationally efficient and easier to implement. Both methods were implemented on simulated, synthetic and real-world data sets of many vegetation indices. The NLS-based method was found to perform better for near-real time beetle infestation detection than the computationally efficient and simpler FIR-based method, and its combination with RGI produced the best performance in terms of accuracy and detection delay. Furthermore, it also showed that deriving test statistics from parameter time series of the the triply modulated cosine function, fitted to the hyper-temporal vegetation index time series using NLS, deals with noise more efficiently and has a clear advantage over an existing method [24] that derives test statistics from the noise time series calculated as a difference between the predicted and observed values.

The MCLT-based method was proposed in Chapter 3. It is the first to use martingale theory in the land cover (beetle infestation) change detection framework. It uses martingale theory to derive test statistics from the μ- parameter time series, instead of log-likelihood ratio based statistics used in the methods explained in Chapter 2. The advantage of using Martingales theory is that the threshold can be selected by using standard Gaussian distribution statistics and does not require tuning based on the labelled training data. Its implementation, for near-real time detection of beetle infestations in pine forests of north America, showed its advantages.

The relative density ratio based change detection method was proposed in Chapter 4. This method derives the μ-parameter time series, of the triply modulated cosine model, using EKF and replaces the log-likelihood ratios based test statistics of the methods proposed in Chapter 2, with RSPRT/CUSUM based test statistics. Furthermore, it improves the RSPRT based test statistics by using the relative density ratios, estimated directly from the data by RULSIF algorithm, as test statistics instead of the traditional likelihood ratios based on Gaussian assumption about the underlying distribution of the data. It also uses an efficient Genetic Algorithm (GA) based threshold tuning for near-real time scenario,

which takes detection delays into account along with the FP and FN, and can be used conveniently in cross validation experiments for model selection. Its implementation on simulated, synthetic and real-world hyper-temporal MODIS beetle infestation data, and comparison with existing methods that use traditional likelihood ratios based on Gaussian assumption about the data distribution, concludes that using assumption-free relative density ratios, estimated directly from the data while deriving RSPRT test statistics, performs better and is more robust to varying data distributions.

The multi-kernel method, for leaf beetle defoliation detection in Eucalyptus plantations of north-eastern Tasmania, Australia using bi-temporal Landsat 7 ETM+ data, was proposed in Chapter 5. It detects changes (defoliation), in the bi-temporal difference image of the raw bands and many vegetation indices, by thresholding the ratios of class posterior probabilities of change and no-change classes for test samples. It models the class posterior probabilities for both the change and no-change classes by using a linear combination of large number of weighted Gaussian kernels (LSPC), and avoids any parametric model assumptions. Furthermore, an automatic model selection procedure was also designed for LSPC, which selects near-optimal model parameters for the change detection framework automatically. The advantage of this model selection procedure is that it does not depend upon the trial and error selection of lists of parameter values by the use as usually done in manual model selection procedure based on 10-fold cross validation, which may lead to suboptimal results. The proposed framework was tested for defoliation detection in eucalyptus nitens in two different test areas, using bi-temporal Landsat 7 ETM+ images. Comparison of its results with the Naive Bayes (NB) and Support Vector Machine (SVM), suggests that non-parametric estimation of the class posterior probabilities has better detection accuracy than the Bayesian classifiers which use parametric model assumption about the data, and is comparable to SVM with added advantages of being capable of handling multi-class problems and assigning probabilistic class-membership to the test

samples. Furthermore, the proposed automated model selection procedure showed improvements in the results of the proposed changed detection framework, over the manual model selection based on 10-fold cross-validation. This non-parametric method is not only limited to remote sensing classification/change detection problems, it can be utilized in other applications as well, in order to replace parametric assumptions about the conditional distribution of the remote sensing data.

The aim of this thesis was to develop efficient remote sensing change detection methods which use freely available coarse or moderate spatial resolution earth observation data to detect insect infestations in natural vegetation. The first three methods were near-real time and the last one is a bi-temporal change detection method. Among the supervised near-real time methods, the relative density ratios based method is an improved and more robust version. The MCLT-based method is the unsupervised version of the NLS-based method in Chapter 2. The scope of the methods proposed in this thesis is not limited to forest cover change detection. They can rather be adapted in a variety of applications in which remote sensing change detection methods have shown promise in recent past, e.g. detection of changes in croplands [31, 32], urban changes [33–35], changes in buildings [36, 37], hydrological changes in soil and coastal areas [38–40], changes in water bodies [41], changes due to floods [42–44], changes caused by fires [45, 46] and deforestation due to logging and human settlements [47–49]. Although, the methods proposed in this thesis showed significant improvements in their respective scopes, like any other method, they also have some limitations which are mentioned in the next section.

6.2 Future Research

In this section, future research ideas based on the findings of this thesis are presented as follows.

- One of the limitations of the methods presented in Chapter 2 and Chapter 4, is that they use a single control charting technique on only the trend parameter μ of a single vegetation index. According to the literature about the statistical control charts, some of them detect small changes quicker, whereas some are quick in detecting bigger changes. Using a hybrid of these control charts, e.g. a combination of CUSUM and $\bar{x}$-chart, may improve the detection accuracy and delays. Furthermore, using it on more than one vegetation indices and their combination with raw bands may generate a variety of features which may improve the performance.

- Another limitation of the methods presented in Chapter 2 and Chapter 4, is that these methods are supervised. As the ground truth data may not be available, the methods, that are independent of labelled data while tuning their thresholds, are needed. Although, the method presented in Chapter 2 does not require threshold tuning, the MCLT implies Gaussian distribution which may be violated in different cases. Therefore, an unsupervised or semi-supervised method which can work in the scenarios when no or partially labelled data is available, and is also free of parametric model assumptions, is required. To this end, one-class-SVM [50] can be used in combination with NLS or EKF and the triply modulated cosine function, to train on the vectors derived only from the no-change (history part) of the time-series. Once the training is done, the one-class-SVM will form a hypersphere enclosing the no-change samples. The query samples can then be declared as no-change or change depending upon whether they lie inside or outside of the hypersphere, respectively.

- The triply modulated cosine model works well in modeling the vegetation index time series. However, it provides only three parameters and is unable to model different scales/resolutions and frequencies. An alternative model can be designed from a linear (weighted) combination of high order Legendre polynomials [51].

Due to orthogonality, Legendre polynomials based model can be truncated after any number of orders because of their inter-polynomials independence in a certain range. This may result in lower prediction error. Furthermore, wavelets multi-resolution analysis and empirical mode decomposition can be used to split the history time-series into different scales. These time-series at different scales can then be predicted separately and merged in a final prediction of the main vegetation index time series. In supervised case, exploiting the fact that certain types of changes may affect certain scales more than the others, and using a combination of only those scales (optimal combination) as feature vectors, may provide better separability between change and no-change classes.

- The Least Squares Probabilisitc Classifier (LSPC) showed a promise in remote sensing change detection (two-class) problem. However, its performance on multi-class remote sensing data and its comparison with with other non-parametric classifiers is still an open research question.

- The LSPC formulation has been tested only with one type of kernel (Gaussian kernel). Driving this formulation using different types of other kernels e.g. linear, polynomial etc, and evaluating its performance on remote sensing data, has not been analysed yet.

- The LSPC based method in Chapter 5 is a supervised change detection technique which estimates class posterior probabilities without any parametric assumption about the data. Adaptation of this method to a remote sensing change detection problem with partially labelled or unlabelled data samples is an open research question.

Bibliography

[1] F. Bovolo and L. Bruzzone, "A split-based approach to unsupervised change detection in large-size multitemporal images: Application to tsunami-damage assessment," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 6, pp. 1658–1670, 2007.

[2] L. Bruzzone and D. F. Prieto, "A minimum-cost thresholding technique for unsupervised change detection," *International Journal of Remote Sensing*, vol. 21, no. 18, pp. 3539–3544, 2000.

[3] T. Celik and K. Ma, "Unsupervised change detection for satellite images using dual-tree complex wavelet transform," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 48, no. 3, pp. 1199–1210, 2010.

[4] P. Coppin, I. Jonckheere, K. Nackaerts, B. Muys, and E. Lambin, "Digital change detection methods in ecosystem monitoring: a review," *International journal of remote sensing*, vol. 25, no. 9, pp. 1565–1596, 2004.

[5] W. Kleynhans, J. C. Olivier, K. J. Wessels, B. P. Salmon, F. van den Bergh, and K. Steenkamp, "Detecting land cover change using an extended kalman filter on modis ndvi time-series data," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, pp. 507–511, May 2011.

[6] E. F. Lambin and A. H. Strahlers, "Change-vector analysis in multitemporal space:

a tool to detect and categorize land-cover change processes using high temporal-resolution satellite data," *Remote Sensing of Environment*, vol. 48, no. 2, pp. 231–244, 1994.

[7] D. Lu, P. Mausel, E. Brondizio, and E. Moran, "Change detection techniques," *International journal of remote sensing*, vol. 25, no. 12, pp. 2365–2401, 2004.

[8] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *International journal of Remote sensing*, vol. 28, no. 5, pp. 823–870, 2007.

[9] B. Salmon, J. Olivier, W. Kleynhans, K. Wessels, F. van den Bergh, and K. Steenkamp, "The use of a multilayer perceptron for detecting new human settlements from a time series of modis images," *International Journal of Applied Earth Observation and Geoinformation*, vol. 13, pp. 873–883, Dec 2011.

[10] B. P. Salmon, J. C. Olivier, K. J. Wessels, W. Kleynhans, F. van den Bergh, and K. C. Steenkamp, "Unsupervised land cover change detection: Meaningful sequential time series analysis," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, pp. 327–335, Jun 2011.

[11] A. Singh, "Review article digital change detection techniques using remotely-sensed data," *International journal of remote sensing*, vol. 10, no. 6, pp. 989–1003, 1989.

[12] S. Franklin, M. Wulder, R. Skakun, and A. Carroll, "Mountain pine beetle red-attack forest damage classification using stratified landsat tm data in british columbia, canada," *Photogrammetric Engineering and Remote Sensing*, vol. 69, pp. 283–288, Mar 2003.

[13] R. H. Fraser and R. Latifovic, "Mapping insect-induced tree defoliation and mortality using coarse spatial resolution satellite imagery," *International Journal of Remote Sensing*, vol. 26, pp. 193–200, Jan 2005.

[14] N. R. Goodwin, N. C. Coops, M. A. Wulder, S. Gillanders, T. A. Schroeder, and T. Nelson, "Estimation of insect infestation dynamics using a temporal sequence of landsat data," *Remote Sensing of Environment*, vol. 112, pp. 3680–3689, Sep 2008.

[15] N. R. Goodwin, S. Magnussen, N. C. Coops, and M. A. Wulder, "Curve fitting of time-series landsat imagery for characterizing a mountain pine beetle infestation," *International Journal of Remote Sensing*, vol. 31, pp. 3263–3271, Jun 2010.

[16] A. J. Meddens, J. A. Hicke, and L. A. Vierling, "Evaluating the potential of multispectral imagery to map multiple stages of tree mortality," *Remote Sensing of Environment*, vol. 115, pp. 1632–1642, Jul 2011.

[17] A. J. Meddens, J. A. Hicke, L. A. Vierling, and A. T. Hudak, "Evaluating methods to detect bark beetle-caused tree mortality using single-date and multi-date landsat imagery," *Remote Sensing of Environment*, vol. 132, pp. 49–58, May 2013.

[18] J. White, M. Wulder, D. Brooks, R. Reich, and R. Wheate, "Detection of red attack stage mountain pine beetle infestation with high spatial resolution satellite imagery," *Remote Sensing of Environment*, vol. 96, pp. 340–351, Jun 2005.

[19] J. White, W. M.A., and D. Grills, "Detecting and mapping mountain pine beetle red-attack damage with spot-5 10-m multispectral imagery," *BC Journal of Ecosystems and Management*, vol. 7, no. 2, pp. 105 – 118, 2006.

[20] J. C. White, N. C. Coops, T. Hilker, M. A. Wulder, and A. L. Carroll, "Detecting mountain pine beetle red attack damage with eo-1 hyperion moisture indices," *International Journal of Remote Sensing*, vol. 28, pp. 2111–2121, May 2007.

[21] M. Wulder, J. White, B. Bentz, M. Alvarez, and N. Coops, "Estimating the probability of mountain pine beetle red-attack damage," *Remote Sensing of Environment*, vol. 101, pp. 150–166, Mar 2006.

[22] M. A. Wulder, C. C. Dymond, J. C. White, D. G. Leckie, and A. L. Carroll, "Sur-

veying mountain pine beetle damage of forests: A review of remote sensing opportunities," *Forest Ecology and Management*, vol. 221, no. 1, pp. 27–41, 2006.

[23] A. Anees and J. Aryal, "Near-real time detection of beetle infestation in pine forests using modis data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 99, pp. 1–11, 2014.

[24] J. Verbesselt and M. Zeileis, A.and Herold, "Near real-time disturbance detection using satellite image time series," *Remote Sensing of Environment*, vol. 123, pp. 98–108, 2012.

[25] T. L. Grobler, E. R. Ackermann, A. J. Van Zyl, J. C. Olivier, W. Kleynhans, and B. P. Salmon, "Using page's cumulative sum test on modis time series to detect land-cover changes," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, pp. 332–336, March 2013.

[26] P. Lombardo and C. Oliver, "Maximum likelihood approach to the detection of changes between multitemporal sar images," *IEE Proceedings on Radar, Sonar and Navigation,*, vol. 148, pp. 200–210, Aug 2001.

[27] J. Otukei and T. Blaschke, "Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms," *International Journal of Applied Earth Observation and Geoinformation*, vol. 12, pp. S27 – S31, February 2010.

[28] G. Yan, J. F. Mas, B. H. P. Maathuis, Z. Xiangmin, and P. M. Van Dijk, "Comparison of pixel-based and object-oriented image classification approaches- a case study in a coal fire area, wuda, inner mongolia, china," *International Journal of Remote Sensing*, vol. 27, no. 18, pp. 4039–4055, 2006.

[29] L. Bruzzone and D. Prieto, "Automatic analysis of the difference image for unsu-

pervised change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, pp. 1171–1182, May 2000.

[30] A. Asmala, "Analysis of maximum likelihood classification on multispectral data," *Applied Mathematical Sciences*, vol. 6, no. 129-132, pp. 6425–6436, 2012.

[31] Z. Zhu and C. E. Woodcock, "Continuous change detection and classification of land cover using all available landsat data," *Remote sensing of Environment*, vol. 144, pp. 152–171, 2014.

[32] J. Dong, X. Xiao, W. Kou, Y. Qin, G. Zhang, L. Li, C. Jin, Y. Zhou, J. Wang, C. Biradar, J. Liu, and B. M. III, "Tracking the dynamics of paddy rice planting area in 1986-2010 through time series landsat images and phenology-based algorithms," *Remote Sensing of Environment*, vol. 160, pp. 99 – 113, 2015.

[33] I. R. Hegazy and M. R. Kaloop, "Monitoring urban growth and land use change detection with {GIS} and remote sensing techniques in daqahlia governorate egypt," *International Journal of Sustainable Built Environment*, vol. 4, no. 1, pp. 117 – 124, 2015.

[34] G. Cao, Y. Li, Y. Liu, and Y. Shang, "Automatic change detection in high-resolution remote-sensing images by means of level set evolution and support vector machine classification," *International Journal of Remote Sensing*, vol. 35, no. 16, pp. 6255–6270, 2014.

[35] B. Wang, S. Choi, Y. Byun, S. Lee, and J. Choi, "Object-based change detection of very high resolution satellite imagery using the cross-sharpening of multitemporal data," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, pp. 1151–1155, May 2015.

[36] X. Huang, L. Zhang, and T. Zhu, "Building change detection from multitemporal high-resolution remotely sensed images based on a morphological building index,"

IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, pp. 105–115, Jan 2014.

[37] S. Nebiker, N. Lack, and M. Deuber, "Building change detection from historical aerial photographs using dense image matching and object-based image analysis," *Remote Sensing*, vol. 6, no. 9, p. 8310, 2014.

[38] K. Rautiainen, T. Parkkinen, J. Lemmetyinen, M. Schwank, A. Wiesmann, J. Ikonen, C. Derksen, S. Davydov, A. Davydova, J. Boike, *et al.*, "Smos prototype algorithm for detecting autumn soil freezing," *Remote Sensing of Environment*, 2016.

[39] G. P. Petropoulos, D. P. Kalivas, H. M. Griffiths, and P. P. Dimou, "Remote sensing and gis analysis for mapping spatio-temporal changes of erosion and deposition of two mediterranean river deltas: The case of the axios and aliakmonas rivers, greece," *International Journal of Applied Earth Observation and Geoinformation*, vol. 35, pp. 217–228, 2015.

[40] B. Tian, Y.-X. Zhou, R. M. Thom, H. L. Diefenderfer, and Q. Yuan, "Detecting wetland changes in shanghai, china using {FORMOSAT} and landsat {TM} imagery," *Journal of Hydrology*, vol. 529, Part 1, pp. 1 – 10, 2015.

[41] K. Rokni, A. Ahmad, A. Selamat, and S. Hazini, "Water feature extraction and change detection using multitemporal landsat imagery," *Remote Sensing*, vol. 6, no. 5, p. 4173, 2014.

[42] D. Lingadevaru, P. Jayakumar, *et al.*, "Flood induced land use and land cover change detection using remote sensing and gis-a study of dhadesugur and siraguppa subwaterheds in lower tungabhadra catchment, karnataka," *International Journal of Geomatics and Geosciences*, vol. 5, no. 4, p. 573, 2015.

[43] V. Klemas, "Remote sensing of floods and flood-prone areas: An overview," *Journal of Coastal Research*, vol. 31, no. 4, pp. 1005–1013, 2014.

[44] Y. Wang, "Advances in remote sensing of flooding," *Water*, vol. 7, no. 11, pp. 6404–6410, 2015.

[45] W. Schroeder, P. Oliva, L. Giglio, B. Quayle, E. Lorenz, and F. Morelli, "Active fire detection using landsat-8/oli data," *Remote Sensing of Environment*, 2015.

[46] J. T. Arnett, N. C. Coops, L. D. Daniels, and R. W. Falls, "Detecting forest damage after a low-severity fire using remote sensing at multiple scales," *International Journal of Applied Earth Observation and Geoinformation*, vol. 35, pp. 239–246, 2015.

[47] W. Kleynhans, B. Salmon, K. Wessels, and J. Olivier, "Rapid detection of new and expanding human settlements in the limpopo province of south africa using a spatio-temporal change detection method," *International Journal of Applied Earth Observation and Geoinformation*, vol. 40, pp. 74 – 80, 2015.

[48] R. B. Mapfumo, A. Murwira, M. Masocha, and R. Andriani, "Detection of subtle deforestation due to logging using satellite remote sensing in wet and dry savanna woodlands of southern africa," *Geocarto International*, vol. 0, no. 0, pp. 1–17, 2016.

[49] J. Reiche, S. de Bruin, D. Hoekman, J. Verbesselt, and M. Herold, "A bayesian approach to combine landsat and alos palsar time series for near real-time deforestation detection," *Remote Sensing*, vol. 7, no. 5, p. 4973, 2015.

[50] B. Schölkopf, J. C. Platt, J. C. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Comput.*, vol. 13, pp. 1443–1471, July 2001.

[51] E. W. Hobson, *The theory of spherical and ellipsoidal harmonics*. Cambridge, University Press, 1931.

Appendix A

Research Papers Emanating From This Thesis:

- A. Anees, J. C. Olivier, M. O'Rielly, and J. Aryal, "Detecting beetle infestations in pine forests using MODIS NDVI time-series data," IEEE, IGARSS, 2013.
- A. Anees and J. Aryal, "Near-real time detection of beetle infestation in pine forests using MODIS data," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, no. 99, pp. 1–11, 2014.
- A. Anees and J. Aryal, "A statistical framework for near-real time detection of beetle infestation in pine forests using MODIS data," Geoscience and Remote Sensing Letters, IEEE, vol. 11, pp. 1717–1721, Oct 2014.
- A. Anees, J. Aryal, M. O'Reilly, and T. Gale, "A relative density ratio-based framework for detection of land cover changes in MODIS NDVI time series," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. PP, no. 99, pp. 1–13, 2015.
- A. Anees, J. Aryal, M. O'Reilly, T. Gale and Tim Wardlaw "Change Detection in Earth Observation Satellite Data using a Robust Multi-Kernel Approach: a case study from Tasmania, Australia," under review in ISPRS Journal of Photogrammetry and Remote Sensing.